EXPLORING GALAXY CLUSTERS AND GROUPS WITH COSMOLOGICAL SIMULATIONS

Nicholas Alistair Henden

Robinson College

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

April 2019

For Mum and Dad.

I made you this.

I don't think you'll enjoy it.

DECLARATION

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed 60,000 word limit.

SUMMARY

Galaxy clusters are the largest gravitationally-bound objects in the Universe, their unparalleled size providing powerful leverage to probe large-scale structure growth and cosmology. At the same time, clusters and groups of galaxies represent unique astrophysical playgrounds in which galaxies interact with each other and with the intervening gas, both of which are strongly influenced by a range of astrophysical processes such as star formation or feedback from supernovae and active galactic nuclei (AGN). In this thesis I present the *Feedback Acting on Baryons in Large-scale Environments* (FABLE) project, a new suite of cosmological hydrodynamical simulations of galaxies, groups and clusters designed to further our understanding of the formation and evolution of these fascinating objects.

Firstly I perform a detailed comparison of the FABLE simulations to low-redshift observations, demonstrating simultaneous agreement with observational constraints on the total stellar and gas mass contents of groups and clusters and the galaxy stellar mass function. I generate synthetic X-ray spectra for the simulated systems and find good agreement with a range of observed X-ray scaling relations. In addition I show that the radial gas profiles of FABLE groups and clusters are a good match to low-redshift observations. Residual deviations in the thermodynamic properties of the cluster core region suggest that more sophisticated AGN feedback modelling or additional physical processes may be needed to explain the observed properties of cluster cores.

Next I extend the analysis of the cluster scaling relations out to high redshift, including a comparison to observational constraints and other simulation predictions. I find that all examined scaling relations deviate from the self-similar prediction in terms of their slope and redshift evolution. These deviations are attributed to a combination of factors, including non-thermal pressure support provided by kinetic motions in the intracluster gas and the effects of non-gravitational physics such as AGN feedback. I also find a significant variation in the scatter about the relations with changing halo mass and redshift, contrary to the assumptions of most observational studies. In addition I investigate the scaling between the Sunyaev-Zel'dovich (SZ) signal and total mass, showing good agreement

with cluster data from *Planck* and the South Pole Telescope over a wide redshift range. I demonstrate the sensitivity of the predicted number of detected clusters in an SZ-selected survey to the assumed SZ scaling relation using several recent observational and simulation constraints.

Finally I investigate the halo mass and redshift dependence of the total baryon contents of FABLE clusters and groups and of the stellar mass, size and shape of brightest cluster galaxies (BCGs). In particular I show that the simulations agree with recent constraints on the (lack of) redshift evolution in the total gas and stellar mass of massive clusters. Furthermore, I use the stellar mass profiles of FABLE BCGs to highlight potential biases in observational studies of BCG growth associated with the assumed light profile and the outer radius of the fit.

ACKNOWLEDGEMENTS

First I would like to thank my supervisors, Ewald Puchwein and Debora Sijacki, for their invaluable direction and sound advice over the past few years. Thank you for humouring my relentless compulsion to write far too many words. Your willingness to proofread my endless waffle deserves a hefty commendation.

This PhD would not have been half as fun without my fellow students, whose company has provided constant entertainment and countless high- (and low-) brow discussions. I don't mind admitting that the prospect of tea, biscuits and a cryptic crossword with the Gridzz Group was often my biggest motivation for arriving at the office before 11am. Of particular note are Aneesh and Tom for so often carrying us to victory and to Lewis for his special brand of crossword solving. The highlight of my week will be no surprise to those who've met the fourth member of office O15 (hint: it's a large pile of boardgames). Many of my fondest memories are of late nights spent assembling doomed spaceships, raising city buildings (twice, for some people) and throwing accusations at my friends (usually the moderator). Special mentions go to Jasleen, Matteo, Lukas, Douglas, Douglas, Adam, Dillon, Sophie and Rob. Sid was also there. I have all of these people and more to thank for my favourite moments in Cambridge; punting in the baking sun and dead of night, impromptu croquet with equipment loosely held together with tape, and formals both delightful and disastrous in nearly equal number, to name but a few. Thanks also to my frequent lunch pals Jasleen, Sophie and Amy for indulging my penchant for pointless facts whether they want to hear them or not, and for occasionally laughing at my jokes. Greenwich tonight? Without a doubt much of the fun I've had during my PhD is thanks to members of the Best Office. Chats and bants with Sophie and Jasleen have ensured that the office always runs at peak inefficiency and maximum entertainment.

I am most thankful to those whose support I have counted on so often during my PhD and over the preceding years. My parents are the kindest people I know and I am stupendously lucky to have them. I owe so many of my achievements, including this one, to the examples they have set me and their unconditional support. And to Katherine I owe my confidence and my (in)sanity. Without the

love and laughter you have brought into my life I would not be half the person I am today.

And finally, See? Told you I would.

CONTENTS

List of Figures

1 INTRODUCTION

Galaxy clusters are big. Really big. As anyone who has read the first line of a research paper on the topic will know, clusters are the largest collapsed objects in the Universe. Having formed from the rarest of tiny density fluctuations in the early Universe, the galaxy clusters that we observe today can weigh more than a thousand trillion suns and consist of hundreds or even thousands of galaxies and huge amounts of hot gas and mysterious dark matter, all bound together by gravity. The downsized counterparts of clusters, called galaxy groups, share many of the same features and, despite containing fewer than about fifty galaxies per group, comprise more than half of all the galaxies in the nearby Universe.

Not to be outdone, galaxy clusters represent a crossroads between cosmology – the study of the evolution of large-scale structure and the Universe as a whole – and astrophysics – the physics and chemistry of astronomical phenomena. Indeed, while their abundance and spatial distribution reflect the overall mass density and initial conditions of the Universe, their deep gravitational potential wells provide a unique astrophysical laboratory in which to study the processes that shape galaxies and impact the hot cluster atmosphere.

Doubtless, galaxy clusters and groups span an astounding range of subject areas. Unfortunately, a full overview of the topic would take up far more space than can fit into a single thesis introduction. Nevertheless, it is my hope that this introduction paints the broad strokes necessary to understand the concepts presented in later chapters.

1.1 A BRIEF HISTORY OF GALAXY CLUSTERS

Despite their name, the galaxy component of galaxy clusters represents only a minute fraction of their total matter content: just ~ 1 per cent of a cluster's mass is bound in stars, while the remainder is approximately 12 per cent gas and 87 per cent invisible dark matter (Chiu et al., 2018). The reason they are not instead called "gas clusters" or "dark clusters" is partly a historical one. Whilst the visible stellar light emitted by galaxies has enabled optical identification of

Figure 1.1: This image of the galaxy cluster Abell 383 is a composite of optical light (blue and white) captured by the Hubble Space Telescope, the Very Large Telescope (VLT) and the Sloan Digital Sky Survey (SDSS) and X-ray data from the Chandra X-ray telescope (purple). The dense collection of galaxies that make up Abell 383 is easy to identify from the optical image. George O. Abell identified this and thousands of other galaxy clusters in the 1950s by visual inspection (with a $3.5\times$ magnifying glass!) of photographic plates from the Palomar Observatory Sky Survey. In contrast to the dense pockets of optical light produced by the galaxies, the X-ray emission of Abell 383 is relatively smooth and extended. These are X-rays emitted by diffuse hot gas trapped within the cluster's gravitational potential well, which by far dominates the mass of stars in galaxies. Image credit: NASA/CXC/Caltech/Newman et al. 2011/Morandi & Limousin 2012 (X-ray); NASA/STScI, ESO/VLT, SDSS (optical).

galaxy clusters as far back as the eighteenth century, the first identification of the hot gas component of clusters via X-ray and microwave instruments occurred much later (in the early 1970s and 1990s, respectively) after this was made possible by technological improvements. In the following I go into more detail about the history of the discovery of galaxy clusters with these different methods and how this has informed our understanding of them.

In optical observations the defining characteristic of galaxy clusters is just that; the clustering of galaxies. In fact, with sufficient resolution of the galaxies, many galaxy clusters can be easily identified by eye from optical images. For example, Fig. 1.1 shows an optical picture centred on the galaxy cluster known as Abell 383. The concentration of galaxies belonging to the cluster clearly stands out from

the surrounding field. Indeed, as far back as 1784 and 1785, Charles Messier and William Herschel first recognised concentrations of galaxies in the constellations of *Virgo* and *Coma Berenices*, which later became known as the Virgo and Coma clusters. However, at the time, it was not known that these were concentrations of galaxies, instead their constituents were referred to simply as 'nebulae'. In part this was because their distances, and therefore their physical sizes, were unknown. It was not until 1925 when Edwin Hubble proved that these nebulae lay at enormous distances that the clusters identified by the likes of Messier and Herschel were confirmed as gigantic, physically associated collections of galaxies (Hubble, 1926).

Only a few years later, Zwicky (1933), Smith (1936) and Zwicky (1937) postulated that the unexpectedly high velocities of galaxies in the Virgo and Coma clusters could only be explained if their gravitational fields were dominated by invisible dark matter. In other words, the masses of these clusters must be far greater than the observed mass in galaxies. The mysterious nature of dark matter coupled with the unparalleled size and mass of galaxy clusters generated much interest in these objects in the following years. This led to the construction of large catalogues of thousands of galaxy clusters, in particular the definitive Abell catalogues (Abell, 1958; Abell et al., 1989), which contain most of the known nearby clusters.

With the advent of X-ray astronomy in the late 1960s it was discovered that galaxy clusters have diffuse, extended X-ray emission due to the presence of hot, ionized gas trapped within their gravitational potential well. Fig. 1.1 shows the X-ray emission of Abell 383 overlaid on the optical image in purple. This image clearly highlights the difference in structure between the diffuse halo gas and the stars condensed into galaxies. Indeed, the vast majority of the gas is not associated with individual galaxies but rather it smoothly fills the space between them, constituting the so-called intracluster medium (ICM). Despite accounting for more than ten times as much mass as all of the stars in a galaxy cluster, the discovery of the ICM accounted for only a small fraction of the missing cluster mass. This finally confirmed the theory of unseen dark matter proposed by Zwicky (1933), which had been disputed in prior decades due to concerns about the methods used to estimate the stellar mass of galaxies. Moreover, the discovery of cluster X-ray emission has proven a great boon to the galaxy cluster community, not only because it facilitates the study of unexplored physics, but also because it provides a means to identify clusters at large cosmological distances (on the order of ten billion light years) in a highly efficient manner (Rosati et al., 2002). I will return to the topic of the X-ray properties of galaxy clusters in Section 1.4.

The hot ICM also leaves an imprint on the sky as observed at microwave wavelengths, which possesses an almost uniform background known as the cosmic microwave background (CMB; see Section 1.2). Sunyaev & Zel'dovich (1970, 1972) predicted that hot gas in clusters would introduce a distortion in the CMB commonly referred to as the Sunyaev-Zel'dovich (SZ) effect. This results from inverse Compton scattering, which occurs when incoming CMB photons are scattered by moving electrons in the hot gas and gain energy as a result. The first high significance detections of galaxy clusters via the SZ effect occurred three decades after its prediction (Birkinshaw, 1999; Carlstrom et al., 2000). The number of SZ-detected clusters has sky-rocketed over the last two decades as instruments capable of measuring the SZ effect have improved and dedicated SZ cluster surveys have been carried out. A particular advantage of the SZ effect is that the SZ signal is nearly independent of distance, unlike optical and X-ray detections, which rely on collecting a sufficient number of photons. As a result, many of these SZ-selected clusters lie at large cosmological distances currently unreachable with other techniques. I summarise the fundamentals of the SZ effect in Section 1.5.

Galaxy clusters sit at the top of a continuous hierarchy of structures that have had time to collapse under their own gravity. While the most massive clusters are more than ten million light years across and contain thousands of galaxies, a much larger fraction of galaxies in the Universe reside in smaller galaxy 'groups' of just two or more galaxies. In many ways galaxy groups represent scaled-down versions of galaxy clusters and most of the concepts described in this introduction apply similarly to groups as to clusters. At the other end of the mass scale, it was assumed for many decades that galaxy clusters were the largest structures in the Universe and were evenly distributed throughout. However, observations carried out since the 1980s have gradually revealed the fact that individual galaxies, groups and clusters are actually interconnected by a network of filaments. Notably, Geller & Huchra (1989) discovered the "Great Wall", an immense filament at least 750 million light years in length and one of the largest known structures in the observable Universe. These galaxy filaments form the boundaries between immense voids where very few galaxies exist, forming a grand structure commonly referred to as the 'cosmic web'. The clearest way to visualise the cosmic web is to make use of very large cosmological simulations to simulate the distribution of dark matter (the dominant mass component of the Universe) in volumes many billions of light years across. An example is shown in Fig. 1.2, which is a visualisation of the dark matter structure of a region of the Millennium-XXL simulation (Angulo et al., 2012). In the next section I

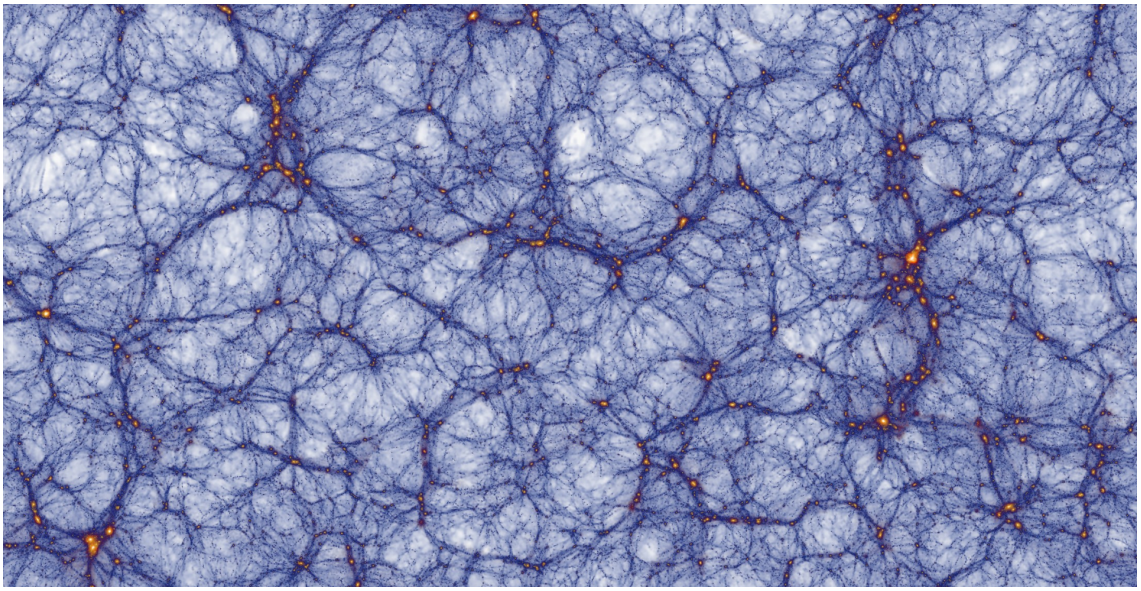


Figure 1.2: The mass density field of dark matter in a region of the Millennium-XXL simulation approximately 740 Mpc (2.4 billion light years) across. The densest regions (galaxy clusters and groups) show up in orange in this colour scale, while the interconnecting filaments and the voids in between them are dark and light blue respectively. The full simulation is a cube with an enormous length of 4.1 Gpc (13.4 billion light years) on a side. This image was obtained using the interactive Millennium-XXL browser.[1] Permission to reproduce this image has been granted by the Max Planck Institute for Astrophysics (MPA). Copyright © MPA/Virgo.

describe our current understanding of how this wealth of structure formed from tiny fluctuations in the very early Universe.

1.2 LARGE-SCALE STRUCTURE FORMATION

The near-uniformity of the CMB – electromagnetic radiation dating back to an early phase of the Universe just 380,000 years after the Big Bang – is in stark contrast to the complexity of structure that we observe today in the form of the cosmic web. This can be appreciated by comparing the highly-concentrated structures in Fig. 1.2 to the map of the CMB shown in Fig. 1.3, which maps the temperature of the CMB across the sky at various different temperature scales. The average temperature of the CMB, which is only 2.725 degrees above absolute zero (i.e. 2.725 degrees Kelvin; symbol K), has been subtracted from the absolute temperature map to highlight the variation in temperature between different points on the sky. At ± 1 K relative to the mean (i.e. 1.725–3.725 K in absolute temperature), the CMB appears completely uniform. It is only by measuring

[1] http://galformod.mpa-garching.mpg.de/mxxlbrowser/

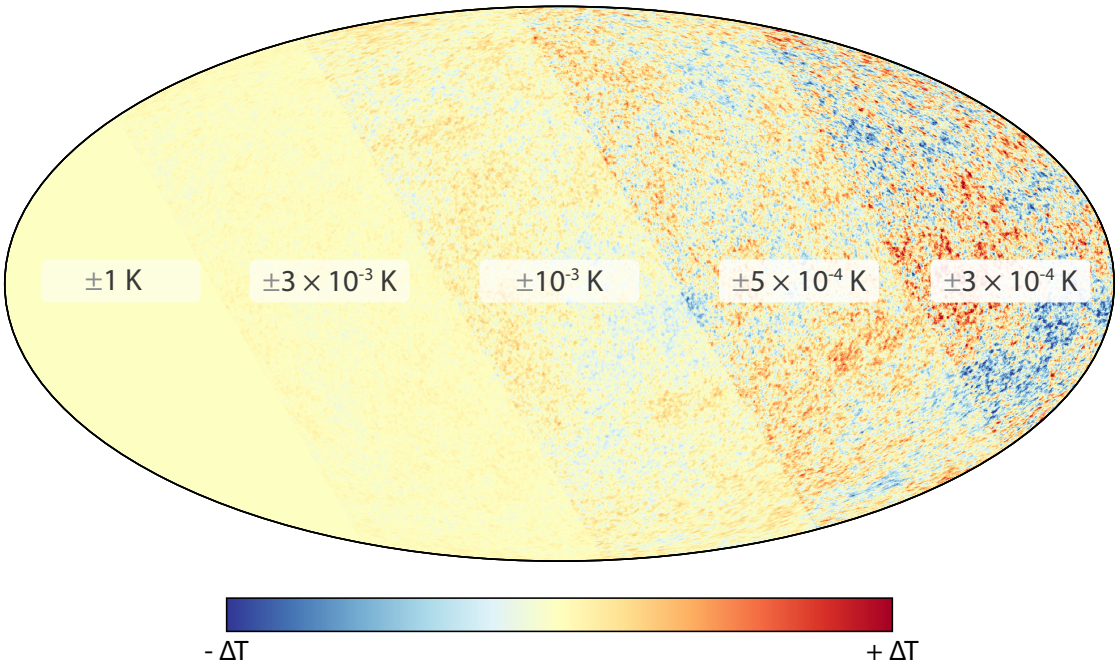


Figure 1.3: An all-sky map of the CMB temperature fluctuations at different temperature scales derived from multi-frequency *Planck* observations.[2] Colour-coding indicates the temperature relative to the mean CMB temperature (2.725 K), with blue and red corresponding to temperatures lower and higher than the mean, respectively. The map is split into five sections corresponding to five different choices for the absolute value of the (symmetric) upper and lower bounds as indicated in the figure. The map is a Mollweide projection of the whole sky in Galactic coordinates, pixelized according to the HEALPix (Gorski et al., 2005) scheme.

the temperature of the CMB to a precision greater than about 1 part in 10^4 that irregularities in the otherwise uniform CMB can be detected. Indeed, by plotting the variation in temperature in smaller and smaller intervals about the mean, Fig. 1.3 shows that the CMB is not perfectly uniform but contains tiny temperature fluctuations of less than one thousandth of a degree.

Over the past 13.8 billion years since the radiation of the CMB was emitted, the density perturbations associated with these temperature fluctuations have grown rapidly under the influence of gravity to form the cosmic web. The gravity of slightly overdense regions attracted matter away from nearby less dense regions, creating galaxies, clusters and filaments on the one hand, and voids on the other. The currently favoured model of structure formation assumes that the matter

[2]The data are obtained from the second *Planck* public data release and constitute the full-mission CMB intensity map at 5 arcminute resolution produced by the COMMANDER method as described in Planck Collaboration IX (2016) and references therein.

component of the Universe is dominated by weakly-interacting cold dark matter (CDM) – the source of invisible mass in galaxy clusters predicted by Zwicky (1933). In this model structures form hierarchically, with small dark matter haloes forming first and merging to create larger structures. I illustrate this process in Section 1.2.3 via a simple toy model in which a cluster grows from a spherical density perturbation in the early Universe.

First, in Section 1.2.1 I describe a recipe for the global geometry and dynamics of the Universe, which forms the basis of the currently favoured model of cosmology, known as ΛCDM. In addition to cold dark matter (of which there is about six times as much mass as ordinary baryonic matter, such as gas and stars), the ΛCDM model describes an accelerating expansion of the Universe driven by an energy component known as 'dark energy' that makes up approximately 70 per cent of the total energy density of the Universe. The nature of dark energy is an area of fervent research in the astrophysical community. Fortunately, in the standard ΛCDM cosmological model its effect on the evolution of the Universe can be described by a single parameter known as the cosmological constant Λ. I introduce Λ in the following section, as well as other parameters of the global cosmological model that define the overall geometry of the Universe, the mean density of its contents, and how its scale changes with time.

1.2.1 *The geometry and dynamics of a homogeneous universe*

According to the Big Bang theory, the Universe was smaller in the past. As one traces backwards in time towards the Big Bang, the constituent gas in the Universe is compressed and becomes hotter and hotter – filling the Universe with hot thermal radiation. Moving the clock forward, the Universe expands and the radiation is stretched to higher wavelengths and lower energies, cooling the gas. Around three minutes after the Big Bang, the temperature is low enough that protons and neutrons can fuse together to create deuterium, helium and lithium nuclei. At this stage, the radiation pressure of the thermal radiation is high enough to prevent any fluctuations in the density of baryonic matter from growing. However, about 380,000 years after the Big Bang, the thermal radiation is cool enough to allow electrons to combine with protons to make neutral hydrogen. This is called the epoch of recombination. In addition, the neutral atoms are transparent to the thermal radiation so that the radiation is able to travel freely through space. This is the radiation we now observe as the CMB. Because the Universe has expanded by a factor of more than one thousand since recombination, the CMB has cooled to an average temperature of just 2.725 K and thus glows

most strongly at microwave wavelengths.

The near uniformity of the CMB implies that the early Universe is isotropic (i.e. the same in all directions). Assuming that we do not occupy a special position in the Universe, this also implies that the early Universe is close to homogeneous (i.e., the same in all locations). In addition, observations show that the distribution of galaxies, galaxy clusters and filaments is consistent with the Universe being, on average, isotropic and homogeneous on large scales (~ 100 Mpc) at the present-day. Consequently, the overall geometry and dynamics of the Universe can be well approximated by that of a fully homogeneous and isotropic one.

1.2.1.1 *The geometry of the Universe*

Assuming isotropy and homogeneity, the (four-dimensional) spacetime geometry of the Universe can be well represented by the relatively simple Robertson-Walker metric, which describes the interval, ds, between two events in spacetime as:

$$ds^2 = c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\right], \tag{1.1}$$

where t is time and r, θ and ϕ are spherical coordinates (radial distance, polar angle and azimuthal angle, respectively) in a 'comoving' coordinate system, which is a set of coordinates that expands along with the Universe. The parameter k describes the curvature of spacetime, which can be negative, positive or zero. Current measurements constrain the spatial curvature to be very close to zero so that the geometry of the Universe appears to be flat.[3] The dimensionless scale factor, $a(t)$, parametrizes the relative expansion of the Universe at some time t relative to its present-day value and is proportional to the mean distance between a pair of objects, e.g. two galaxy clusters, moving with the expansion of the Universe. By definition, $a(t_0) = 1$, where t_0 is the present age of the Universe.

As mentioned above, electromagnetic radiation is stretched by the expansion of the Universe so that, by the time it reaches Earth, the wavelength of the observed light ($\lambda_{\rm obs}$) is longer than when it was emitted ($\lambda_{\rm em}$). This phenomenon is known as redshifting, as the wavelength of light is shifted towards the redder end of the spectrum (i.e. longer wavelengths). Conveniently, measuring the change in wavelength provides a way to determine the scale factor of the Universe when the radiation was emitted. This is called the cosmological redshift, z, and it is defined as:

$$z = \frac{1}{a} - 1 = \frac{\lambda_{\rm obs} - \lambda_{\rm em}}{\lambda_{\rm em}}. \tag{1.2}$$

[3]This means, for example, that the angles in a triangle (regardless of its size) add up to precisely 180 degrees and not less than (negative curvature) or more than 180 (positive curvature).

Observationally, the determination of the cosmological redshift requires accounting for the redshifting of light caused by the peculiar motion of the emitting object relative to the expansion of the Universe.

1.2.1.2 *The dynamics of the Universe*

To determine the time evolution of $a(t)$ (i.e., the expansion rate of the Universe) requires the Einstein field equations from Albert Einstein's general theory of relativity, which describes gravity as the result of spacetime being curved by mass and energy. The Robertson-Walker metric of equation 1.1 is a solution of the Einstein field equations provided that the Friedmann equations are satisfied. The Friedmann equations form the basis of the ΛCDM cosmological model and can be written as:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3}\,, \tag{1.3}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3}\,, \tag{1.4}$$

where $a = a(t)$, $\rho c^2 = \rho(t)c^2$ is the mean energy density of the Universe, $p = p(t)$ is the pressure owing to that energy density, and Λ is the cosmological constant.[4] Equation 1.4, which describes the acceleration of universal expansion, shows that the energy density and pressure cause the expansion to *decelerate*, while a positive cosmological constant causes an opposite *acceleration*. In particular, a positive cosmological constant acts like an additional energy density equal to $\frac{\Lambda c^2}{8\pi G}$ that exerts a negative pressure, thereby accelerating the expansion. Observations have shown that the expansion of our Universe is accelerating, which implies a positive value for Λ that, in the ΛCDM model, is assumed to come from the repulsive gravity of dark energy (although see e.g. Frieman et al. 2008 for a discussion of the various theoretical ideas proposed to explain the observed acceleration of the Universe, which may or may not invoke dark energy and/or a cosmological constant).

Equation 1.3 is greatly simplified in a flat universe (i.e. $k = 0$) for which the energy density is predominantly in the form of matter with zero dark energy (i.e., $\Lambda = 0$ with ρ equal to the average matter density). Rearranging equation 1.3 with these assumptions yields an expression for the density of such a universe, called the critical density:

$$\rho_{\rm crit} = \frac{3H^2}{8\pi G}, \tag{1.5}$$

[4]Note that, with the Friedmann equations defined as in equations 1.3 and 1.4, ρc is the mean energy density of the Universe without a contribution from the cosmological constant.

where the Hubble parameter, H, is equal to the quantity $\frac{\dot{a}}{a}$ and describes the expansion rate of the universe at a given point in time. For example, the present-day value of H, called H_0, is constrained by observations of the CMB to be $H_0 \simeq 67.8\ \mathrm{km\,s^{-1}Mpc^{-1}}$ (Planck Collaboration XIII, 2016). This means that a galaxy at a distance of 1 Mpc (about 3 million light-years) is moving away from us at $67.8\ \mathrm{km\,s^{-1}}$ due to the expansion of the Universe. The critical density is a useful quantity because it determines the average energy density required for a universe to possess a flat geometry. This is true even for a universe with a non-zero cosmological constant, such as in $\Lambda\mathrm{CDM}$, although in this case one must consider the average energy density of the universe *including* the energy density owing to the cosmological constant (the dark energy density in $\Lambda\mathrm{CDM}$).

Because the expansion of our own Universe is accelerating, its expansion rate, as described by H, is not the same today as it was in the past. Thus, to describe the full dynamics of the Universe requires an expression for H as a function of time (or, equivalently, the scale factor), which can be derived from the Friedmann equations defined above. From a practical perspective, this expression would be most useful in terms of quantities that can be observed today. This requires an understanding of how the average energy density of the Universe depends on the scale factor. The density is typically separated into three known forms of energy: matter, radiation and dark energy, all of which can influence the overall expansion history. In the standard $\Lambda\mathrm{CDM}$ model, dark energy is characterised by the cosmological constant and thus its energy density remains constant as the Universe expands. The radiation density in the Universe became subdominant to the other forms of energy about 47,000 years after the Big Bang (e.g. it is only a factor $\sim 10^{-4}$ of the total energy density today) and can be neglected for simplicity. As a result, ρ in equations 1.3 and 1.4 can be approximated as the density of (non-relativistic) matter, $\rho \approx \rho_m$. Because the total amount of matter stays the same as the Universe expands, ρ_m must decrease with time proportional to the inverse volume of the Universe. Hence, ρ_m varies with the scale factor and redshift as $\rho_m = \rho_{m,0}\, a^{-3} = \rho_{m,0}\,(1+z)^3$, where $\rho_{m,0}$ is the density of matter today. It is common to write $\rho_{m,0}$ as a fraction of the present-day value of the critical density ($\rho_{\mathrm{crit},0}$):

$$\Omega_m \equiv \frac{\rho_{m,0}}{\rho_{\mathrm{crit},0}} = \frac{8\pi G}{3H_0^2}\rho_{m,0}, \tag{1.6}$$

where Ω_m is called the matter density parameter. Current constraints place Ω_m at a value of ~ 0.31 (Planck Collaboration XIII, 2016). Since the total energy density of the Universe (matter, radiation and dark energy) is close to the critical density, this implies that 31 per cent of the energy in the Universe is in the form of matter.

The matter density parameter corresponding just to baryons (i.e. not including dark matter) is called Ω_b and is constrained to be ~ 0.05. Hence, there is about five times more dark matter than baryons. Similarly, the energy density corresponding to the cosmological constant normalised by the critical density gives the dark energy density parameter:

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H_0^2}. \tag{1.7}$$

Current constraints find $\Omega_\Lambda \simeq 0.69$ (Planck Collaboration XIII, 2016). Hence, dark energy constitutes about 69 per cent of the energy density of the Universe. Putting this all together, and writing the scale factor in terms of redshift (equation 1.2), the Hubble parameter, $H(z)$, at some redshift z can be written as:

$$E(z) \equiv \frac{H(z)}{H_0} = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}. \tag{1.8}$$

where $E(z)$ is the Hubble parameter normalised to its present-day value, H_0. Note that this equation holds only for a universe that is spatially flat and for which the radiation density is a negligible fraction of the critical density. The parameter $E(z)$ will become particularly useful in later chapters where it is used frequently to describe the evolution of a number of cluster properties.

1.2.2 *Density fluctuations in the early Universe*

Strictly speaking, a universe described exactly by the Robertson-Walker metric and the Friedmann equations contains no galaxy clusters, galaxies or stars, as this would violate the assumption of homogeneity and isotropy. Fortunately, this relatively simple model is a good approximation to the evolution of the real Universe on very large scales, and models that account for the lumpiness of the Universe can be added onto it as extensions. Indeed, the ΛCDM model contains a set of parameters that specify the initial spectrum of density perturbations in the (otherwise homogeneous) early Universe that grew into the galaxies and clusters that we see today. As an example, the scalar spectral index, n_s, is used to describe how the density fluctuations vary with scale. The physical meaning of n_s is best understood in terms of how the amplitude of the primordial fluctuations varies on different mass scales. By considering the variance in mass within identical spherical volumes, it can be shown that the typical mass fluctuation scales with the mass M as (e.g. Voit et al. 2005; Mo et al. 2010):

$$\frac{\delta M}{M} \propto M^{-(n_s+3)/6}. \tag{1.9}$$

This equation implies that current constraints on n_s, which are close to unity ($n_s \simeq 0.967$; Planck Collaboration XIII 2016), mean that the primordial density

perturbations have larger amplitudes on smaller mass scales.[5] In fact, this is true for cold dark matter models in general, which have $n_s > -3$ over all length scales (e.g. Blumenthal et al. 1984). As a result, smaller mass fluctuations will, on average, begin to collapse earlier than larger mass fluctuations. In this case, structure formation proceeds in a hierarchical fashion, with less massive haloes forming prior to more massive ones and eventually merging to create even larger structures. An example of this process is given in the following section, which describes a simple model for the growth and collapse of a cluster-scale halo in a cold dark matter universe.

1.2.3 *The growth of a cluster: The spherical collapse model*

The initial linear growth of primordial density perturbations can be described analytically via perturbation theory (e.g. Heath 1977). However, the linear approximation eventually breaks down once the density fluctuation is comparable in magnitude to the background density. Further evolution involves highly non-linear interactions in which small haloes merge and coalesce to form larger and larger structures. Over the past few decades, a number of approximate analytic models have been developed to study the non-linear evolution of dark matter perturbations (e.g. Gunn & Gott 1972; Fillmore & Goldreich 1984; Bertschinger 1985; Nusser & Sheth 1999; Ascasibar et al. 2004; Williams et al. 2004; Del Popolo 2012). The complexity of these models has increased over time by the gradual addition of ever more realistic dynamics of the growth process, such as the influence of baryonic collapse on the dark matter distribution (e.g. Ryden & Gunn 1987; Blumenthal et al. 1986; Gnedin et al. 2004). A full description of these models is beyond the scope of this introduction, however many of the basic features of cluster formation can be illustrated with a simple model based on the standard spherical collapse model introduced in Gunn & Gott (1972).

First consider, in a spatially flat, matter-dominated universe with $\Omega_\Lambda = 0$, a uniform spherical region that has an excess density $\delta\rho$ above the average density in the universe, ρ. Initially the sphere will expand at the same rate as the rest of the universe. However, the extra gravitational attraction of its excess mass will cause its own expansion to be slower than the universe as a whole. The fractional

[5] Technically, equation 1.9, with the power-law index set to n_s, applies only on very large scales (on the order of $\gtrsim 100$ comoving Mpc). On smaller scales, the growth of perturbations is suppressed during the radiation-dominated era due to the effects of radiation pressure. This produces a weaker dependence of $\frac{\delta M}{M}$ on the mass scale M, resulting in a $\frac{\delta M}{M}$ that is roughly constant on the smallest scales. More specifically, if n_s in equation 1.9 is replaced with the 'effective' spectral index n for a given mass or length scale, then $n = n_s \approx 1$ on $\gtrsim 100$ Mpc scales but drops to $n \approx -2.5$ on the scale of low-mass galaxies.

density enhancement, $\delta\rho/\rho$, will thus become larger and larger until eventually the sphere stops expanding and begins to collapse under the influence of gravity. The sphere reaches a maximum radius at this point called the turn-around radius, $r_{\rm max}$. In fact, it can be shown that turn-around occurs when the sphere is approximately 5.55 times as dense as the background universe (e.g. Mo et al. 2010).

Assuming collisionless dark matter, a uniform sphere will eventually collapse to a point of infinite density in a time T_c that scales with the size of the initial density perturbation approximately as $T_c \propto (\delta\rho/\rho)^{-3/2}$ (Gunn & Gott, 1972). Clearly this did not occur in the real Universe. In part this is because the primordial density perturbations that eventually form clusters are not uniform in density. Instead there are density fluctuations on all scales, with smaller mass scale perturbations superposed on top of the larger one that form less massive haloes such as galaxies. Because the initial density enhancement, $\delta\rho/\rho$, is larger on smaller mass scales (Section 1.2.2), the galaxies will have a shorter collapse time. Thus, by the time the cluster has stopped expanding, individual galaxies have already formed. Because the galaxies are not distributed symmetrically, they will miss each other as they fall through the cluster and expand back out again. This process occurs repeatedly, during which gravitational interactions between galaxies produce a time-varying gravitational potential that randomises their velocities in a process known as "violent relaxation" (Lynden-Bell, 1967), leading to a state of virial equilibrium.

The relative density of the cluster at virialisation can be approximated by noting that the total energetics should satisfy the virial theorem for a system in equilibrium:

$$E_k = -\frac{1}{2}E_p, \tag{1.10}$$

where E_k is the total kinetic energy per unit mass and E_p is the total gravitational potential energy per unit mass. At turn-around, the cluster is momentarily still and so the total energy ($E = E_k + E_p$) is equal to its potential energy: $E = E_p = -GM/r_{\rm max}$, where M is the total enclosed mass within $r_{\rm max}$. As the sphere collapses, some of this potential energy is converted into kinetic energy but the total energy and total mass are conserved. It follows that the virial theorem is satisfied when the cluster has collapsed to $r_{\rm max}/2$, since at this time:

$$E_k = E - E_p = -GM/r_{\rm max} + GM/(r_{\rm max}/2) = GM/r_{\rm max} = -\frac{1}{2}E_p. \tag{1.11}$$

After virialisation the cluster is therefore eight times denser than it was at turn-around. In this time, the background universe has continued to expand and is less dense by a factor of four.[6] Since the cluster was 5.55 times denser than the

[6]This assumes that the time taken to reach virial equilibrium is equal to the time taken to

background universe at turn-around, the virialised cluster is now over-dense by a factor of approximately $5.55 \times 8 \times 4 \approx 178$. A more rigorous treatment (e.g. Peebles 1980) similarly yields $18\pi^2 \approx 178$, while more general formulae (e.g. for a universe with a non-zero cosmological constant) can be found in Bryan & Norman (1998). Results such as these have motivated astronomers to define the virial radius of a cluster as the spherical radius in which the average matter density is 200 times the critical density (see Section 1.6). By comparison, the density contrast predicted by linear perturbation theory at the collapse time of the uniform spherical perturbation is approximately 1.69 (e.g. Mo et al. 2010). This highlights just how fast structures can grow when the perturbation enters the non-linear regime.

The spherical collapse model provides a useful insight into the initial non-linear growth of clusters. For example, Press & Schechter (1974) applied the model to a Gaussian initial density field to create the useful Press-Schecter formalism, which allows one to estimate the *mass function* of collapsed objects (i.e. their abundance as a function of mass). Of course this simple model misses some other important characteristics of cluster formation that have been included in later analytic models. For example, if the initial over-density is triaxial (as all galaxy clusters are observed to be) rather than spherically symmetric, then the perturbation will tend to collapse first along the shortest axis to form a disc, then along the intermediate axis to form a filament and eventually along the longest axis to form a cluster (e.g. Lin et al. 1965; White & Silk 1979; Sheth et al. 2001; Shen et al. 2006). In addition, the initial perturbation will not have a well-defined edge, but rather an extended density profile out to large radii. The shape of this profile determines the rate at which mass accretes onto the cluster (e.g. Fillmore & Goldreich 1984; Bertschinger 1985; Hoffman & Shaham 1985) and, in combination with the process of relaxation, the density profile of the cluster after virialisation (e.g. Nusser & Sheth 1999; Manrique et al. 2003).

Analytical approximations such as these provide useful guidelines for the time-scale of halo collapse and have been used to great effect in developing approximate statistical models for the evolution of large halo populations such as the Press-Schecter formalism. However, all such models rely on strong simplifying assumptions. Ultimately, a full treatment of hierarchical structure formation requires numerical cosmological simulations that follow the growth and collapse of realistic density perturbations by solving the equations of motion for discrete

reach turn-around. Taking $\rho \propto a^{-3}$, $k = 0$ and $\Lambda = 0$ for a flat, matter-dominated universe in equation 1.3, it can be shown that the scale factor scales with time as $a \propto t^{2/3}$ and the average density as $\rho \propto t^{-2}$. Thus, a factor 2 increase in time corresponds to a factor 4 decrease in density.

particles of dark matter. I will return to this topic in Section 1.8.

1.2.4 *Clusters as cosmological probes*

The growth of galaxy clusters from the rarest primordial density perturbations makes them sensitive probes of cosmology (see e.g. Allen et al. 2011 for a review). In particular, the number density of clusters as a function of their mass can be used to constrain the cosmological parameters Ω_m (equation 1.6) and σ_8, which describes the amplitude of linear fluctuations smoothed on a scale of $8\,h^{-1}$ Mpc. In addition, the redshift dependence of the cluster counts can be used to constrain the equation of state of dark energy (see e.g. Weinberg et al. 2013).

The critical ingredient required to constrain cosmology using cluster counts is a theoretical prediction for the expected number of clusters as a function of mass, redshift and cosmology (e.g. from numerical simulations; see Section 1.8). This can be used to calculate the likelihood that the observed cluster counts would be realised in a universe with a particular cosmology and enables a determination of the most probable cosmological parameters based on the observations.

This process is complicated by the fact that cluster mass is not a direct observable. As a result, the abundance of clusters as a function of mass must be inferred from the abundance as a function of some observable (e.g. X-ray temperature or SZ signal) using a mass–observable scaling relation. Biases in the calibration of this relation can introduce significant systematic uncertainties on the cosmological parameters derived from cluster counts (e.g. Allen et al. 2011). In fact, for recent cluster surveys these often outweigh the statistical uncertainties (e.g. Planck Collaboration XXIV 2016).

1.3 GALAXIES

Like clusters, observations suggest that galaxies are dominated by dark matter (e.g. Roberts & Rots 1973; Einasto et al. 1974; Ostriker et al. 1974; Rubin et al. 1978; Tyson et al. 1984). This motivated White & Rees (1978) to propose a two-stage theory for galaxy formation in which dark matter haloes form through hierarchical merging and then, bound in the gravitational potential well of the dark matter, baryonic gas cools, condenses and forms luminous galaxies with a variety of properties (e.g. Efstathiou & Silk 1983; Blumenthal et al. 1984).

1.3.1 *Galaxies in* ΛCDM

The gravitational collapse of a density perturbation creates strong shocks in the gas, which raises its temperature. The gas can also lower its temperature via radiative cooling, which can occur via a number of excitation and de-excitation mechanisms depending on the temperature, density and chemical composition of the gas (see e.g. Mo et al. 2010). If cooling is slow – for example, if the gas is not very dense – the system relaxes into hydrostatic equilibrium within the gravitational potential well of the dark matter halo as the outward force of gas pressure balances the inward gravitational force. If radiative cooling is efficient (which may occur in the denser inner regions after hydrostatic equilibrium has been reached), the gas loses pressure support and accumulates as dense, cold gas at the centre of the dark matter halo, forming a protogalaxy.

Eventually the density of gas in the protogalaxy will exceed that of the dark matter in the centre of the halo and become self-gravitating. Continued efficient cooling can cause the formation of dense, cold gas clouds in the protogalaxy within which star formation can occur. Additional mergers trigger even more star formation, eventually giving rise to a visible galaxy. Unfortunately the precise details of these processes are not yet fully understood. One issue in particular that remains unclear is the mass distribution of newly-formed stars, called the initial mass function (IMF). The IMF is particularly important because the masses of these stars largely determine how they are expected to evolve.

1.3.2 *Feedback processes*

Early models of galaxy formation in a cold dark matter Universe found that radiative cooling in protogalaxies should be very efficient (e.g. Binney 1977; Rees & Ostriker 1977; White & Rees 1978; Thoul & Weinberg 1995). The gas should therefore condense, form stars and, due to the loss of pressure support, allow more gas to flow onto the protogalaxy in a continual process of rapid star formation (e.g. White & Frenk 1991; Katz et al. 1992; Kauffmann et al. 1993; Katz et al. 1996). However, this is at odds with observations, which show that only a small fraction of baryons are in the form of cold gas or stars. Part of the puzzle can be explained by photoionization of the gas by massive stars or by the ultra-violet (UV) background radiation produced by star-forming galaxies and black holes at high-redshift. Photoionization can both reduce the cooling rate of the gas and also raise its temperature, thereby suppressing ('quenching') star formation, particularly in low-mass galaxies (Efstathiou, 1992; Bullock et al., 2000; Dijkstra

et al., 2004; Okamoto et al., 2008). However, photoionization alone is not sufficient to explain the observations, which implies that an additional form of so-called 'feedback' must be heating the gas.

1.3.2.1 *Stellar feedback*

One solution to this problem lies with feedback from high-mass stars in the form of radiation, stellar winds and supernovae. In terms of their impact on galaxy-scale properties, the most influential of these stellar feedback mechanisms comes from supernovae, a class of exploding stars that can produce enormous amounts of energy. The radiation and blast waves from these supernovae can heat up the gas that constitutes the interstellar medium (ISM) of the galaxy, driving galactic outflows that expel potentially star-forming gas from the galaxy and slowing further infall of gas from its environment. Supernovae feedback can impact a variety of galaxy properties, including morphology, stellar motions, the metal content of the gas and, perhaps most importantly, the overall star formation rate and total stellar mass. Indeed, a number of theoretical models predict that stellar feedback may explain the discrepancy between the large numbers of low-mass galaxies predicted in a cold dark matter universe and the much lower number of observed low-mass galaxies (e.g. White & Frenk 1991; Somerville & Primack 1999; Springel & Hernquist 2003b; Murray et al. 2005; Crain et al. 2009; Schaye et al. 2010; Guo et al. 2011). On the other hand, in the most massive galaxies quenching by stellar feedback is ineffective because the energy it transfers into the ISM is insufficient to overcome the rapid cooling of the central gas, which has been compressed to high densities by the deep gravitational potential wells of these systems (e.g. McCarthy et al. 2012; Kannan et al. 2014).

1.3.2.2 *AGN feedback*

A more energetic quenching mechanism is provided by active galactic nuclei (AGN), which are supermassive black holes (SMBH) in an active phase of accretion. SMBHs are present at the centre of almost all massive galaxies and are capable of releasing vast amounts of energy into the ISM of the galaxy as AGN. If this energy couples efficiently to the gas it can significantly suppress star formation within the galaxy. The effects of AGN feedback can even extend far beyond the galaxy itself, as seen in X-ray observations of galaxy cluster cores (see Section 1.4). This is especially remarkable given that the ratio of the size of a SMBH to its host galaxy is so small (on the order of $10^{-8} - 10^{-9}$). Similar to the way stellar feedback can suppress star formation in low-mass haloes, AGN feedback is often invoked in analytical models and simulations to explain the discrepancy between the mass

function of galaxies predicted in cold dark matter models and the observed galaxy stellar mass function at the high mass end (e.g. Bower et al. 2006; Croton et al. 2006; Schaye et al. 2010).

AGN feedback is thought to act in one of two modes, deemed the radio-mode and the quasar-mode (e.g. Churazov et al. 2005; Sijacki & Springel 2006; Weigel et al. 2017). The radio-mode of feedback is associated with the production of jets by moderately accreting black holes. Observational signatures of this mode were first identified in X-ray observations of the ICM in the form of jet-inflated bubbles originating from the central cluster galaxy. I will describe the radio-mode in more detail in Section 1.4, but for the purpose of this section it is sufficient to say that radio-mode feedback, like quasar-mode feedback, is thought to keep the gas surrounding galaxies warm and prevent it from condensing onto the galaxy.

The quasar-mode, also called the radiative or wind mode, is more difficult to observe directly because the AGN couples to its surroundings via emitted radiation, most of which is absorbed by material within the galaxy (see e.g. King & Pounds 2015 for a review). It is thought that most of the radiation originates from an accretion disk, which pushes matter away along the disk axis, however it is currently unclear precisely how the radiation interacts with the surrounding gas. The most plausible theories include a high-velocity wind (which may be driven by radiation) or direct radiation pressure on surrounding dust. The primary observational signature of quasar-mode AGN feedback is a fast outflow of gas from the host galaxy. Many galactic outflows have been detected in various gas phases, such as molecular, atomic and ionised gas (e.g. Greene et al. 2011; Cano-Díaz et al. 2012; Liu et al. 2013; Veilleux et al. 2013; Harrison et al. 2014; Rupke et al. 2017). Bubbles of hot gas coincident with such winds have also been detected in X-rays (Greene et al., 2014; Sartori et al., 2016; Lansbury et al., 2018) and potentially via the SZ effect (Lacy et al., 2019).

The quasar-mode is thought to operate when the AGN is accreting rapidly and was probably most effective at redshift $z \sim 2$–3 when observations show galaxies were most gas-rich and black hole accretion rates were high. In fact, the evolution of the luminosity density of AGN, which tracks the average black hole accretion rate, roughly follows the evolution of the average star formation rate in the Universe quite well (e.g. Boyle & Terlevich 1998; Silverman et al. 2008; Aird et al. 2010; Madau & Dickinson 2014; Aird et al. 2015, 2018). Furthermore, quasar-mode AGN activity is believed to explain observed correlations between central black hole mass and properties of the host galaxy, such as the stellar velocity dispersion or galactic bulge mass (e.g. Haring & Rix 2004; Kormendy & Ho 2013; McConnell & Ma 2013). These results suggest that galaxies and SMBHs are

intrinsically linked during their evolution, with feedback from accreting SMBHs regulating the growth of their host galaxies. Direct observational evidence for quasar-mode feedback suppressing star formation in galaxies is still patchy, but highly suggestive correlations do exist (e.g. Farrah et al. 2012; Wylezalek & Zakamska 2016; Baron et al. 2017). Perhaps the most convincing argument for AGN feedback comes from semi-analytic models and simulations, which typically require some form of AGN feedback in order to match observations of the stellar mass build-up in massive galaxies (e.g. Di Matteo et al. 2005; Sijacki & Springel 2006; Sijacki et al. 2007; Booth & Schaye 2009; Vogelsberger et al. 2014a; Schaye et al. 2015).

1.3.3 *Galaxies in clusters*

Galaxy clusters consist of a large dark matter halo hosting a massive central galaxy and many smaller satellite galaxies. The central galaxy, typically called the brightest cluster galaxy (BCG) due to its dominant mass and luminosity compared with the other cluster galaxies, is thought to form mainly via galaxy mergers. This can occur either through major mergers, in which two merging systems have similar mass, or minor mergers, in which one system is significantly less massive than the other (typically less than $1/3$). The former can occur as part of a merger between the main cluster halo and another large halo of comparable mass, in which case violent relaxation causes the central galaxies to merge and the satellites to enter a quasi-equilibrium state. Build-up of central galaxy mass via minor mergers usually occurs as orbiting satellites transfer energy to the main cluster halo by dynamical friction, causing them to spiral inwards and eventually merge with the central galaxy. Interactions between the infalling satellite and the cluster's gravitational potential or other satellite galaxies cause tidal effects that remove mass from the outer regions of the satellite. Furthermore, as the satellite approaches the central regions, its gas component experiences a ram pressure with the increasingly dense ICM, which can strip gas from the galaxy. This enriches the ICM with metals from the ISM that are produced within the stars of the galaxy and recycled into the ISM in a process called chemical enrichment.

BCGs often possess a very diffuse and extended outer envelope of stellar light that accounts for a significant fraction of the total stellar light in the cluster. Part of this emission may be considered 'intracluster light' (ICL), which is associated with stars that are not gravitationally bound to any particular galaxy. It is thought that the ICL consists mainly of stars stripped from satellite galaxies by tidal interactions (e.g. Contini et al. 2014; Montes & Trujillo 2018; DeMaio et al. 2018), although some

contribution may result from violent mergers with the central galaxy (e.g. Conroy et al. 2007; Murante et al. 2007) or in-situ star formation (e.g. Puchwein et al. 2010; Tonnesen & Bryan 2012). Some recent observations estimate that the ICL consists of 5–20 per cent of the total mass of stars in clusters (e.g. Burke et al. 2015; Montes & Trujillo 2018; Jiménez-Teja et al. 2018). On the other hand, a number of previous studies have measured significantly higher ICL fractions of 20–50 per cent (e.g. Seigar et al. 2007; Zibetti 2007; McGee & Balogh 2010; Gonzalez et al. 2013). The latter are generally in better agreement with simulation predictions (e.g. Sommer-Larsen et al. 2005; Purcell et al. 2007; Puchwein et al. 2010; Cui et al. 2014a; Pillepich et al. 2018b), although some studies find low ICL fractions of $\approx$ 10–20 per cent (e.g. Willman et al. 2004; Rudick et al. 2011). These discrepancies are likely due to a combination of factors. From an observational point of view, the amount of light in the ICL is difficult to constrain due to its very low surface brightness. In addition, because the extended profile of the BCG smoothly blends into the ICL, distinguishing one component from the other is somewhat ambiguous, both observationally (e.g. Rudick et al. 2011; Bernardi et al. 2013) and in simulations (e.g. Puchwein et al. 2010; Cui et al. 2014a).

1.4 CLUSTERS IN X-RAYS

As clusters collapse under their own gravity and grow via mergers with smaller systems, the gas within their potential wells (consisting mostly of hydrogen and helium) is heated by adiabatic compression and shocks to X-ray emitting temperatures of several keV (tens of million of degrees). The two primary mechanisms for this X-ray emission are thermal bremsstrahlung and line emission. Bremsstrahlung photons are produced when free electrons in the hot gas are decelerated by electromagnetic interactions with ions, losing kinetic energy and converting it into radiation. The intensity of the radiation is weakly temperature-dependent (scaling like $\sim T^{1/2}$) but strongly dependent on the gas density (scaling like $\sim \rho_{\rm gas}^2$).[7] Line emission is produced by elements such as iron that, heavily ionized, emit X-rays in a narrow frequency band as electrons drop from one energy level to a lower one. The X-ray emission of the ICM is a complicated combination of these and other processes. Fortunately, X-ray spectra can be generated with relative ease with software packages such as the publicly available `XSPEC`, which employs X-ray emission models that take all of these factors into account.

[7]Strictly speaking, the intensity is proportional to the product of the number densities of electrons and ions in the plasma summed over all species of ions (e.g. Rybicki & Lightman 1979).

1.4.1 *Evidence for radio-mode AGN feedback*

In 1991, X-ray observations of the central galaxy in the Perseus cluster revealed clear depressions in the X-ray surface brightness to the north and south of the galaxy (Böhringer et al., 1993). After the launch of the *Chandra* X-ray observatory in 1999, similar cavities in the X-ray emission were seen in the majority of X-ray bright clusters undergoing efficient radiative cooling in their cores (e.g. Dunn & Fabian 2006). Comparison with radio images of the cluster cores showed that the innermost cavities usually coincide with radio lobes produced by a central AGN (e.g. Böhringer et al. 1993; Dunn et al. 2005; Dunn & Fabian 2008). It was realised that the efficient radiative cooling in these systems has caused gas to condense onto the central galaxy, some of which is accreted by the AGN. The accreting AGN generates powerful twin jets that inflate almost-spherical bubbles of relativistic plasma in the surrounding ICM (e.g. Binney & Tabor 1995; Omma et al. 2004; McNamara et al. 2005). These bubbles rise buoyantly in the hot gas atmosphere, stretching and expanding until they are no longer visible with current instrumentation. It is expected that the bubbles transfer some of their energy to the ICM during this process, although the precise mechanism and its efficiency are currently under debate (see e.g. Yang & Reynolds 2016b and references therein). In this way, AGN jet-inflated bubbles act as a form of feedback, raising the entropy of the gas and slowing the process of cooling, condensation and star formation. This so-called radio-mode of feedback, which is thought to be active at low AGN accretion rates, thus acts similarly to the quasar-mode of AGN feedback described in Section 1.3.2.2, which operates at high accretion rates.

AGN feedback is often invoked as a way to explain the lack of so-called 'cooling flows' in observed clusters (see e.g. McNamara & Nulsen 2007 for a full discussion). The concept of a cooling flow stems from the realisation that the strong X-ray emission of the ICM in the cluster core should correspond to a rapid cooling of the central gas, which should lower pressure support in the cluster centre and cause gas from larger radii to flow inwards (the cooling flow). Yet there is very little evidence for cool X-ray emitting gas in clusters (e.g. Peterson et al. 2001, 2003; Tamura et al. 2001, 2003; Sanders et al. 2008; Hudson et al. 2010) and significantly less cold molecular gas than expected from pure cooling models ($\lesssim$ 10 per cent; e.g. Edge 2001; Edge & Frayer 2003; Salomé & Combes 2003; Mittal et al. 2011; Russell et al. 2014, 2017; Pulido et al. 2018). Furthermore, the increased density of cold gas in the central regions should have led to runaway star formation, which is not observed (e.g. Johnstone et al. 1987; O'Dea et al. 2008; Rawle et al. 2012; Mittal et al. 2015; Cooke et al. 2016; McDonald et al. 2018). This

implies the need for a source of thermal energy in cluster cores that is capable of heating the gas to sufficient temperature to prevent such rapid cooling. A number of theories have been put forward to explain the lack of cooling flows in clusters, such as thermal conduction from hot gas at larger radii (e.g. Ruszkowski & Begelman 2002; Conroy & Ostriker 2008; Ruszkowski & Oh 2010) or stirring of gas by the motions of substructures (e.g. Fujita et al. 2004; Ruszkowski & Oh 2011), however AGN feedback is the most popular explanation due to the amount of energy AGN can theoretically emit over their lifetimes, as well as measurements of the power associated with observed X-ray cavities (see Fabian 2012 for a review). Moreover, hydrodynamical simulations have consistently shown that models which mimic the inflation of jet cavities by injecting hot bubbles (e.g. Churazov et al. 2001; Quilis et al. 2001; Sijacki & Springel 2006; Sijacki et al. 2007, 2015), as well as models that attempt to model the inflation of the jet cavities themselves (e.g. Bourne & Sijacki 2017; Weinberger et al. 2017b; Bourne et al. 2019), are able to offset overcooling, effectively disrupting cooling flows in clusters and significantly reducing star formation rates in massive galaxies.

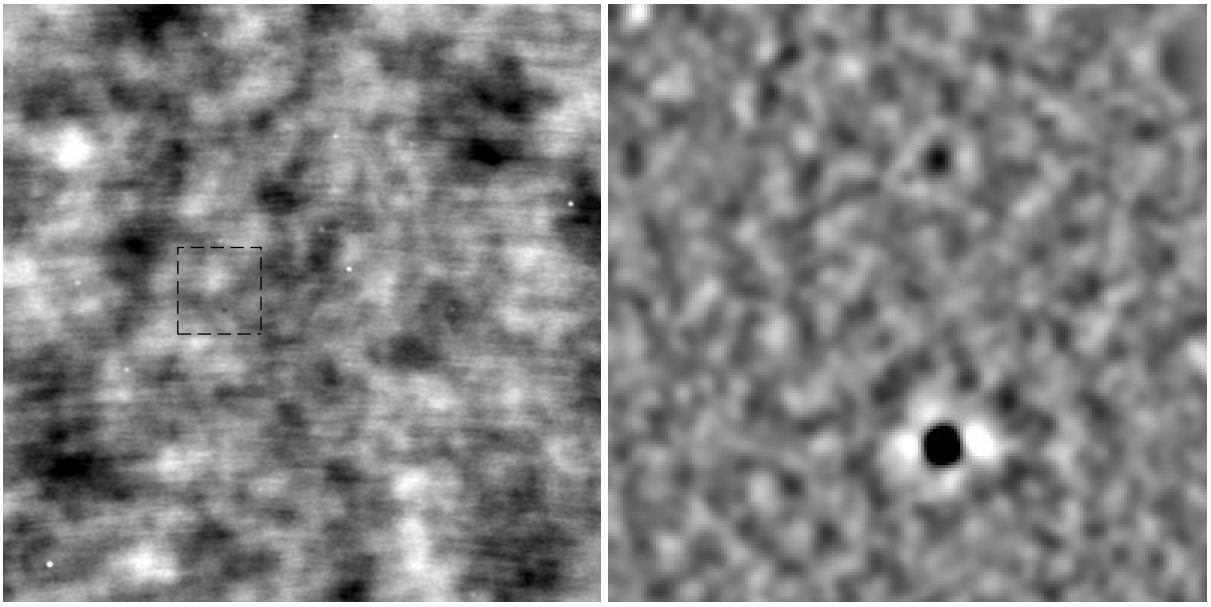
1.5 CLUSTERS IN SZ

The SZ effect is usually divided into two parts: a *thermal* effect due to the random thermal motions of hot electrons, and a *kinetic* effect due to bulk motion of the gas with respect to the CMB. Only the thermal SZ effect will be considered in this thesis, which dominates over the kinetic SZ effect for galaxy clusters.

The thermal SZ effect arises as CMB photons are scattered upwards in energy by their encounters with energetic electrons in a process called inverse Compton scattering. This causes a change in the observed spectral flux density of the CMB that is dependent on the energy (or, equivalently, the frequency) of the CMB photons being observed. For example, at frequencies greater than about 217 GHz the spectral flux density in the direction of a galaxy cluster will increase relative to the average, whereas at lower frequencies it will decrease. This occurs because the total photon number of photons is conserved and thus the number of relatively low energy photons is depleted as these are scattered towards higher energies. An illustration of how the SZ signal can be used to identify galaxy clusters in maps of the CMB is shown in Fig. 1.4.

At fixed frequency the amplitude of the intensity change in the CMB due to the thermal SZ effect scales with the Compton y parameter (e.g. Birkinshaw 1999),

$$y = \int \frac{k_B T_e}{m_e c^2} n_e \sigma_T \, dl, \tag{1.12}$$



(a) 150 GHz map cutout (b) Cluster-filtered map, zoomed in to 1°-by-1°

Figure 1.4: Panel (a) shows a 6°-by-6° cut-out of the 150 GHz map of the CMB from SPT-SZ data (Bleem et al., 2015) over the temperature range $\pm 3 \times 10^{-4}$ K. The primary CMB anisotropies on degree scales are clearly visible. The white spots are emissive point sources, typically radio or infra-red galaxies. At this observing frequency, the SZ signal is negative and therefore clusters appear as 'shadows' against the CMB. Panel (b) zooms in on the 1°-by-1° area near one such cluster as indicated by the dashed square in panel (a). A spatial filter has been applied to enhance the signal-to-noise ratio (SNR) on cluster scales and the map is in units of SNR over the range $-5 < \mathrm{SNR} < 5$. The cluster appearing as a strong SZ decrement in the bottom of panel (b) has been optically confirmed at a redshift of $z = 1.132$. This figure is reproduced from panels (b) and (d) in fig. 1 of Bleem et al. (2015). Copyright © AAS. Reproduced with permission.

where the integration follows the line-of-sight, T_e and n_e are the electron temperature and electron number density of the scattering medium, σ_T is the Thomson cross section and k_B, c and m_e are fundamental constants (the Boltzmann constant, speed of light and electron rest mass, respectively). Since the product of temperature and density is proportional to pressure, the Compton y parameter is effectively a measure of the integrated electron pressure of the hot gas along the line-of-sight. One important result is that the Compton y parameter is independent of redshift and depends only on the properties of the scattering medium (in this case, the ICM). This enables the detection of clusters at high redshift that may be too faint to identify at X-ray and optical wavelengths.

1.6 WEIGHING GALAXY CLUSTERS

Quantifying the total bound mass in galaxy clusters and groups is an important endeavour from both a cosmological and an astrophysical viewpoint. For example, complex physical processes that govern galaxy formation and the heating/cooling of the ICM can cause baryonic observables to follow the dark matter imperfectly. Consequently, a comparison between observables and the total mass distribution can provide valuable information concerning these poorly understood processes. From a cosmological perspective, the number density of groups and clusters as a function of their mass is sensitive to the underlying cosmology and can be used to derive constraints on the cosmological parameters (see Section 1.2.4).

The most commonly-used definition of cluster mass – and the one used throughout this thesis – is the total mass, M_{Δ}, enclosed within a sphere of radius r_{Δ} inside which the average mass density is Δ times the critical mass density of the Universe, $\rho_{\rm crit}(z)$, at the redshift, z, of the cluster. The mass is thus

$$M_{\Delta} = \frac{4\pi}{3} \Delta \, \rho_{\rm crit}(z) \, r_{\Delta}^{3} \, . \tag{1.13}$$

As discussed in Section 1.2.3, a value of $\Delta = 180$ corresponds approximately to the virial radius of a galaxy cluster in a matter-dominated universe. Indeed, a value of $\Delta = 200$ is often used to define the virial radius of a cluster (although, for a more realistic universe with $\Omega_{\Lambda} = 0.3$, the virial density contrast is $\Delta \approx 100$; Bryan & Norman 1998). In practice, observers often prefer a higher threshold of $\Delta = 500$ because the corresponding radius, r_{500}, is more accessible to the current generation of X-ray telescopes. Values of the corresponding mass, M_{500}, span the range $\sim 10^{13} - 10^{14}$ solar masses ($M_{\odot}$) for galaxy groups and $\sim 10^{14} \, M_{\odot}$ to a few times $10^{15} \, M_{\odot}$ for galaxy clusters. The separation between groups and clusters at $M_{500} \sim 10^{14} \, M_{\odot}$ is not distinct, but it provides a useful point of division when discussing the different scales occupied by these two populations (e.g. $r_{500} \approx 600$ kpc for a typical group with $M_{500} = 5 \times 10^{13} \, M_{\odot}$ and $r_{500} \approx 1300$ kpc for a typical cluster with $M_{500} = 6 \times 10^{14} \, M_{\odot}$).

A number of different methods can be used to estimate the total bound mass distribution of cluster and groups. The two most commonly-used techniques are X-ray hydrostatic analyses and gravitational lensing. As discussed in the following, cluster masses estimated via these two methods can differ significantly. As many of the results presented in later chapters rely on comparisons to such masses, I dedicate much of the following discussion to the possible sources of the discrepancies between them.

1.6.1 *X-ray Hydrostatic Masses*

The density and temperature distribution of the ICM as inferred from its X-ray emission can be directly related to the gravitational potential of the system under the assumption that the gas is in hydrostatic equilibrium within the cluster potential (i.e. the pressure gradient is balanced by the gradient of local gravitational potential). Combined with the ideal gas equation of state and assuming spherical symmetry it can be shown that:

$$\frac{d\ln\rho_g(r)}{d\ln r} + \frac{d\ln T(r)}{d\ln r} = -\frac{G\mu m_p}{k_B}\frac{M(<r)}{r\,T(r)} \tag{1.14}$$

where $\rho_g(r)$ and $T(r)$ are the radial gas density and temperature profiles and $M(<r)$ is the total mass enclosed within a radius r. Here G is the gravitational constant, μ is the mean molecular weight and m_p is the rest mass of a proton.

The fundamental assumption behind this technique is that the cluster is in hydrostatic equilibrium. However, growing theoretical and observational evidence argues that this assumption is often violated, most likely due to random gas motions (i.e. turbulence), bulk flows or gas acceleration that alter the state of hydrostatic equilibrium (e.g. Kay et al. 2004; Mahdavi et al. 2008; Lau et al. 2009, 2013; Smith et al. 2016; Biffi et al. 2016; Hurier & Angulo 2018). These kinetic motions can provide non-thermal pressure support to partially balance the gravitational field, which causes the hydrostatic mass to underestimate the true mass.

Indeed, numerical simulations are essentially unanimous in predicting that X-ray hydrostatic masses underestimate the true halo mass. They do not however agree on the size of the so-called X-ray hydrostatic mass bias, which typically has some dependence on the mass, redshift or dynamical state of the cluster (e.g. Rasia et al. 2006; Nagai et al. 2007a; Piffaretti & Valdarnini 2008; Rasia et al. 2012, 2014; Barnes et al. 2017b; Henson et al. 2017). Most simulation studies predict that turbulence is the dominant contributor to the hydrostatic mass bias, independent of the dynamical state of the cluster, while bulk motions and gas acceleration are important in merging clusters (e.g. Rasia et al. 2006; Meneghetti et al. 2010; Lau et al. 2013; Nelson et al. 2014a,b; Shi et al. 2015, 2016). Aside from a violation of hydrostatic equilibrium, X-ray hydrostatic masses may also be biased by other aspects of cluster physics, such as gas inhomogeneities (e.g. Rasia et al. 2012; Khedekar et al. 2013), or by observational biases, such as spectroscopic temperature bias (e.g. Rasia et al. 2006; Nagai et al. 2007a; Piffaretti & Valdarnini 2008; Rasia et al. 2012) or instrument calibration (e.g. Donahue et al. 2014; Israel et al. 2015; Schellenberger et al. 2015).

Observationally, a large effort has been made in recent years to compare X-ray hydrostatic masses to masses derived via gravitational lensing that, in contrast to hydrostatic masses, are insensitive to the dynamical state of the cluster (see Section 1.6.2). Unfortunately, the degree to which X-ray hydrostatic masses are biased low compared to lensing masses is inconsistent between studies, with values ranging from zero to ~ 30 per cent (e.g. Richard et al. 2010; Mahdavi et al. 2013; Donahue et al. 2014; Melin & Bartlett 2015; Sereno & Ettori 2015a; Smith et al. 2016; Applegate et al. 2016; Battaglia et al. 2016; Hurier & Angulo 2018; Parroni et al. 2017; Sereno et al. 2017; Penna-Lima et al. 2017; Medezinski et al. 2018). This effort was partly motivated by the results of cosmological studies of cluster samples such as the *Planck* SZ sample, for which the error budget was dominated by uncertainties in the mass calibration based on X-ray hydrostatic masses. Even assuming an X-ray hydrostatic mass bias of 20 per cent as motivated by numerical simulations, Planck Collaboration XX (2014) found a significant tension between cosmological constraints from *Planck* cluster counts and the *Planck* primary CMB constraints. In the most recent cosmological analysis of *Planck* data, Planck Collaboration VI (2018) find that consistency between these different constraints requires that X-ray hydrostatic masses are biased low by 38 ± 3 per cent. This lies on the upper end of the bias measurements estimated from weak lensing studies, although it is consistent with most recent studies to within 1-sigma (see discussion in Planck Collaboration VI 2018). In addition, Zubeldia & Challinor (2019) have recently revisited the mass calibration of *Planck* clusters with updated mass measurements from gravitational lensing signatures in the CMB, finding a mass bias of 29 ± 10 per cent that is consistent with Planck Collaboration VI (2018). On the other hand, a large X-ray mass bias (~ 40 per cent) as seemingly required by *Planck* primary CMB constraints implies a level of gas depletion in massive clusters greatly exceeding that predicted by cosmological simulations with reasonable feedback prescriptions (see e.g. Eckert et al. 2016), which poses a major challenge to our current understanding of galaxy clusters.

Fortunately, in the near future the increasing number of clusters with lensing mass estimates will help to narrow down the true level of X-ray mass bias (see Section 1.6.2). Furthermore, CMB lensing constraints with upcoming experiments such as Advanced ACTpol (Henderson et al., 2016) and SPT-3G (Benson et al., 2014) are expected to improve the mass uncertainty to approximately 3 per cent (Pratt et al., 2019) and to sub-percent accuracy with a CMB Stage-4 experiment (Louis & Alonso, 2017). Similarly, a large amount of effort continues to be put into archival and follow-up studies using the *Chandra* and *XMM-Newton* X-ray telescopes, including SZ-selected cluster samples (e.g. Lovisari et al. 2017; Reiprich

2017; Bulbul et al. 2019). Such studies will enable a fair comparison between large samples of clusters that were previously hampered by the heterogeneity of the data and concerns surrounding the representativeness of X-ray selected samples (e.g. Rossetti et al. 2016, 2017; Andrade-Santos et al. 2017). In the future, high-resolution X-ray spectroscopy from upcoming X-ray missions such as XRISM and Athena will enable measurements of the non-thermal pressure due to bulk and random motions in the ICM, thereby correcting for the hydrostatic mass bias (Ota et al., 2018; Roncarelli et al., 2018).

1.6.2 *Gravitational Lensing*

Gravitational lensing describes the phenomena resulting from the bending of light rays by gravitational fields. The observed lensing effects in clusters are classified as either 'strong' or 'weak' lensing. Strong lensing is characterised by multiple images or giant arcs. Several examples of the latter can be seen near the centre of the cluster Abell 383 in the optical image shown in Fig. 1.1. A limitation of strong lensing for the purpose of estimating cluster masses is that it occurs only occasionally in rich, centrally-concentrated clusters of galaxies and, moreover, it typically allows a mass determination only of the inner region of a lensing cluster where the projected mass density is highest. On the other hand, weak lensing in the form of weakly distorted background galaxies is present in every cluster and can be analysed in a statistical sense to constrain cluster mass density profiles out to large radii. For these reasons, weak lensing is a much more common method than strong lensing for estimating total cluster masses.

In essence, measuring the total mass distribution of a galaxy cluster via weak lensing involves measuring the shapes of a large number of background galaxies and comparing the statistical average with the expected value for an isotropic distribution of shapes. As this thesis does not deal directly with the determination of cluster masses via weak lensing, I defer to Bartelmann & Schneider (2001) for a detailed explanation of the theory and to Mandelbaum (2018) for a review of the practical aspects of weak lensing analyses in current and near-future surveys.

Gravitational lensing is arguably the most direct probe of the total mass distribution in galaxy clusters among existing techniques. Furthermore, it is relatively insensitive to the dynamical state of the cluster. This makes weak lensing masses a particularly useful check of other total mass estimates, such as X-ray hydrostatic masses. Importantly for our analysis in Chapters 2 and 3, weak lensing is generally considered the most direct technique for measuring the absolute calibration of the relations between cluster observables and total mass.

However, weak lensing mass measurements are not free from systematics themselves and typically possess larger scatter than X-ray hydrostatic masses. In part this is because the lensing signal describes all of the mass along the line-of-sight from the lens to the source. This can introduce significant bias and scatter in the recovered weak lensing masses due to the presence of large-scale structure in projection (e.g. Marian et al. 2010; Becker & Kravtsov 2011; Hoekstra et al. 2011). In addition, cluster-sized haloes are often not spherical but triaxial in shape, so that the weak lensing mass can over- or underestimate the true mass depending on whether the major axis is aligned with, or transverse to, the line of sight (e.g. Oguri et al. 2005; Corless & King 2007; Meneghetti et al. 2010; Bahé et al. 2012). Non-negligible biases and additional scatter can also be introduced by issues such as miscentring (e.g. George et al. 2012), cluster substructure (e.g. Meneghetti et al. 2010; Bahé et al. 2012; Henson et al. 2017), or contamination by unlensed cluster and foreground galaxies (e.g. Medezinski et al. 2010; Okabe & Smith 2016).

As the shape distortion of galaxies via weak lensing is so small compared with the intrinsic, randomly-oriented galaxy shapes (often called "shape noise"), precise weak lensing mass measurements require a very large ensemble of background galaxies. In the past, the deep imaging required to obtain a sufficient number density of faint background galaxies has limited the number of weak lensing studies and essentially prevented large-scale weak lensing surveys. However, several such surveys have started in recent years, such as the Kilo-Degree survey (de Jong et al., 2013), the Dark Energy Survey (Dark Energy Survey Collaboration, 2016) and the Hyper Suprime-Cam survey (Aihara et al., 2018). In the coming decade, surveys such as Euclid (Laureijs et al., 2011), the Large Synoptic Survey Telescope (LSST; LSST Dark Energy Science Collaboration 2012) and the Wide Field Infrared Survey Telescope (WFIRST; Spergel et al. 2015) will further increase the breadth and precision of weak lensing measurements.

1.7 SELF-SIMILARITY

A variety of optical, X-ray and SZ cluster observables correlate strongly with the total mass of the cluster. Indeed, one of the most common methods to estimate masses for a large sample of clusters is from the relationships between cluster mass and readily-observable cluster properties such as X-ray luminosity or temperature. These mass–observable scaling relations can often be approximated as simple power laws (see e.g. the relation between SZ signal and total mass in Fig. 2.10). This power law form is expected from simplified models in which galaxy clusters are simply scaled versions of each other, as in the widely-used self-similar model

developed by Kaiser (1986). The self-similar model assumes that both gravity and the initial density perturbations have no preferred scale and that no additional scales are introduced into the problem by the physical processes that shape the properties of forming clusters. In this model, cluster formation is characterised by a single collapse epoch (as in the spherical collapse model described in Section 1.2.3) and the ICM is heated only by the shocks associated with the collapse. Despite its simplicity, the self-similar model provides predictions for the power law slopes of the scaling relations that are remarkably similar to observational results for massive clusters (e.g. Mantz et al. 2010; Maughan et al. 2012; Mantz et al. 2016a). Self-similarity thus provides a useful baseline for studying the cluster scaling relations. In particular, comparing the observed properties of clusters to the self-similar expectation can help to isolate the impact of physical processes such as radiative cooling and feedback during cluster formation.

The self-similar model predicts that the power law slopes of the scaling relations should not evolve with redshift. However, evolution is expected in the normalisations of the relations, the form of which is dependent on the mass definition. If the total mass of a cluster, M_Δ, is defined via equation 1.13 to be the mass enclosing a spherical region of radius r_Δ with an average density Δ times the critical density of the Universe, $\rho_{\rm crit}$ (equation 1.5), then the redshift evolution of the self-similar scaling relations is due only to the evolving critical density of the Universe. With this definition, the expected self-similar slope and redshift evolution of the scaling relations can be straightforwardly derived.

Combining the mass definition in equation 1.13 and the critical density defined in 1.5 leads to the useful scaling between cluster size and mass

$$r_\Delta \propto M_\Delta^{1/3} E(z)^{-2/3} \,. \tag{1.15}$$

The self-similar model assumes that the ICM is in equilibrium within the cluster gravitational potential so that the total cluster mass is given by the hydrostatic equilibrium equation (equation 1.14) evaluated at $r = r_\Delta$. Thus, M_Δ is proportional to $r_\Delta T(r_\Delta)$ and the sum of the logarithmic slopes of the gas density and temperature profiles at r_Δ. Another assumption made by Kaiser (1986) is that the density and temperature profiles themselves are self-similar. In this case the logarithmic slopes of the profiles are independent of mass and it can be shown that $T(r_\Delta)$ is proportional to T_Δ, the average temperature within r_Δ (Kravtsov & Borgani, 2012). The average temperature is therefore related to the total mass via

$$T_\Delta \propto \frac{M_\Delta}{r_\Delta} \propto E(z)^{2/3} M_\Delta^{2/3} \,, \tag{1.16}$$

where r_Δ has been substituted from equation 1.15.

Assuming that the bolometric X-ray emission of a cluster is dominated by thermal bremsstrahlung, the bolometric X-ray luminosity can be approximated as (e.g. Sarazin 1986)

$$L_{\Delta}^{\rm bol} \propto \rho_g^2 \, r_{\Delta}^3 \, \Lambda(T) \propto \frac{M_{\rm gas,\Delta}^2}{r_{\Delta}^3} \, T^{1/2} \tag{1.17}$$

where $M_{\rm gas,\Delta} \equiv \rho_g r_{\Delta}^3$ is the total gas mass and $\Lambda(T) \propto T^{1/2}$ is the cooling function for bolometric emission. Since the gas density profile is assumed to scale self-similarly, the total gas mass scales as (Kravtsov & Borgani, 2012)

$$M_{\rm gas,\Delta} \propto M_{\Delta} \, . \tag{1.18}$$

The bolometric X-ray luminosity is thus related to mass and temperature via

$$L_{\Delta}^{\rm bol} \propto \frac{M_{\Delta}^2}{r_{\Delta}^3} \, T_{\Delta}^{1/2} \propto E(z)^2 \, M_{\Delta} \, T_{\Delta}^{1/2} \tag{1.19}$$

where r_{Δ} has been substituted from equation 1.15. Substituting for temperature or mass from equation 1.16 gives the self-similar scaling relations

$$L_{\Delta}^{\rm bol} \propto E(z)^{7/3} M_{\Delta}^{4/3} \tag{1.20}$$

and

$$L_{\Delta}^{\rm bol} \propto E(z) \, T_{\Delta}^2 \, . \tag{1.21}$$

Lastly, the integrated SZ signal, $Y_{\rm SZ,\Delta}$, and its X-ray analogue, $Y_{\rm X}$, are related to the gas mass and temperature as $Y_{\rm SZ,\Delta} \propto Y_{\rm X,\Delta} \equiv M_{\rm gas,\Delta} \, T_{\Delta}$. Substituting for temperature from equation 1.16 and relating the gas mass to the total mass with equation 1.18 gives the self-similar relations

$$Y_{\rm SZ,\Delta} \propto Y_{\rm X,\Delta} \propto E(z)^{2/3} \, M_{\Delta}^{5/3} \, . \tag{1.22}$$

1.8 NUMERICAL SIMULATIONS

As described in Section 1.2, the highly non-linear processes of collapse that drive the growth of large-scale structure can only be approximated analytically with numerous simplifying assumptions. Instead a far more accurate method is to perform three-dimensional 'N-body' simulations of dark matter. Indeed, the FABLE simulations,[8] which are the focus of later chapters, make use of one such simulation (Millennium-XXL; Angulo et al. 2012) to choose initial conditions from which we can be sure to grow suitably massive galaxy groups and clusters. In

[8]The acronym FABLE stands for *Feedback Acting on Baryons in Large-scale Environments*.

Section 1.8.1 I describe some of the N-body techniques used to model dark matter in simulations.

In addition to dark matter, the formation of galaxies and galaxy clusters also relies heavily on interactions between baryons and a wide variety of astrophysical processes (see Section 1.3). Two major approaches have been developed to circumvent this problem. One approach is the semi-analytic method, which uses knowledge about the structure and assembly history of cold dark matter haloes from an N-body simulation to model the galaxies that form within them. I cover this approach briefly in Section 1.8.2. A different approach is to use hydrodynamical simulations such as FABLE, which numerically solve the non-linear equations governing dark matter, baryons and galaxy formation physics simultaneously. These are discussed in Sections 1.8.3 and 1.8.4.

1.8.1 *N-body simulations*

Since the early 1990s a large number of studies have performed cosmological N-body simulations of cold dark matter to study various properties of dark matter haloes, such as their density profiles (e.g. Navarro et al. 1997, 2004; Graham et al. 2006), triaxiality (e.g. Frenk et al. 1988; Cole & Lacey 1996; Jing & Suto 2002; Allgood et al. 2006), angular momentum distributions (e.g. Warren et al. 1992; Vitvitska et al. 2002) and merger histories (e.g. Lacey & Cole 1993; Fakhouri et al. 2010). One of the most important uses for N-body simulations today is their ability to predict the mass function of collapsed objects beyond the precision of analytic approximations such as the Press-Schecter formalism, including their redshift and cosmology dependence (e.g. Jenkins et al. 2001; Tinker et al. 2008; Crocce et al. 2010; Courtin et al. 2010). Excitingly, these calibrations form the basis for the use of galaxy clusters as tools to constrain cosmological models (see Section 1.2.4).

1.8.1.1 *Gravity solvers*

The fundamental requirement for the simulation of large-scale structure formation is the ability to calculate the force of gravity at a given point in an extended mass distribution. To achieve this numerically, a common approach is to discretise the dark matter into a large number of point-particles and to calculate the gravitational force on each particle due to all others.[9] The size and direction of

[9] In practice a softening of the gravitational force is introduced to limit the importance of close encounters between particles. This is necessary because the simulation particles are many orders of magnitude more massive than any physical counterpart and therefore the gravitational attraction between two point-particles at small spatial separation can become unphysically large. The standard solution to this problem is to smooth the mass of each particle over a finite volume (usually a sphere), the radius of which is determined by the chosen gravitational softening length.

the force determines the acceleration of each particle and thus their positions and velocities after some finite time step. This procedure is iterated until the desired time interval is reached. The smaller the chosen time step, the more accurate the result, but at the cost of extra computation time.

The simplest and most exact way to obtain the gravitational force on each particle is to measure the distance between all pairs of particles, calculate the gravitational force between each pair (which is inversely proportional to the square of their separation) and finally sum the forces on each particle. Unfortunately, the computation time required for this direct summation approach scales very poorly with the total number of particles. For example, doubling the number of particles in the simulation would quadruple the time taken for completion – in general, the computation time scales like N^2, where N is the number of particles.

Various numerical techniques have been developed to reduce the computational cost with manageable losses in accuracy, most notably particle-mesh (PM) methods and tree algorithms. The PM method (e.g. Hockney & Eastwood 1981) interpolates the particle masses onto a regular grid and computes the gravitational potential via Fourier transform. This approach is fast even for large numbers of particles, however the force resolution is limited by the resolution of the mesh, which can lead to degeneracies in highly-clustered regions. Tree algorithms are used to restructure the particles into a hierarchical 'tree', which enables groups of particles that are sufficiently far-away to be treated as if they were a single mass. This speeds up the computation time, which scales with particle number as $N \log N$. A common choice is the hierarchical oct-tree algorithm of Barnes & Hut (1986), which recursively divides the simulation volume into eight cubic cells until no cell contains more than one particle. The gravity solver in AREPO (the simulation code employed in this thesis) is a combination of these two techniques (called the TreePM method), which splits the gravity calculation into short- and long-range components calculated via an oct-tree algorithm and PM method, respectively. The TreePM method was used for example to simulate the large-scale matter distribution shown in Fig. 1.2.

1.8.1.2 *Cosmological initial conditions*

Cosmological simulations begin from initial conditions that mimic the state of the Universe at very high redshift (typically $z \sim 100$). To create these initial conditions one must first assume a background cosmological model, such as ΛCDM, that determines the geometry, dynamics and composition of the Universe and the primordial spectrum of density perturbations (see Section 1.2.1 and 1.2.2). Consequently, the standard technique is to assume an initial Gaussian random density

field with statistical properties specified by the chosen cosmological model.[10] The perturbations are then evolved to the desired redshift according to linear growth theory and a so-called transfer function, which describes scale-imprinting effects from different physical processes such as pressure gradients (e.g. Bardeen et al. 1986; Holtzman 1989; Eisenstein & Hu 1998, 1999). To obtain the initial positions and velocities of the simulation particles, the standard approach is to position equal-mass particles on a uniform grid and to use the perturbation spectrum and the Zeldovich (1970) approximation to compute their initial displacements and velocities relative to the grid (Efstathiou et al., 1985). Unfortunately, using a uniform grid can introduce an artificial pattern in the particle distribution that persists visibly until virialisation. To avoid this problem it is common to arrange the particles in a "glass-like" configuration before displacement, which is achieved by first placing the particles in random positions and then moving them by small amounts in the opposite direction of gravity until they "freeze" in comoving coordinates (Baugh et al., 1995).

1.8.1.3 *Zoom-in simulations*

The most massive dark matter haloes, such as those belonging to galaxy clusters, form only from the rarest peaks in the initial matter density field. Thus, to simulate an appreciable number of galaxy clusters requires the simulation of a very large volume. Unfortunately, this requirement severely limits the resolution of the simulation. This makes it difficult to resolve relatively small objects, such as galaxies, at the same time as more massive systems, like galaxy clusters. One way around this problem is the use of 'zoom-in' simulations (e.g. Katz & White 1993; Power et al. 2003; Navarro et al. 2004).

In cosmological zoom-in simulations, a region of interest from a large-volume N-body simulation is re-simulated at high resolution while the neighbourhood of that region is left at a coarser resolution. This means that the desired object can be followed at high resolution whilst taking into account realistic tidal torques generated by the low resolution neighbourhood, which are required for realistic structure formation. To simulate a galaxy cluster via this technique the first step is to select all of the dark matter particles that make up the cluster at the endpoint of the simulation (e.g. all of the particles out to some multiple of the virial radius). The target particles are then traced back to the cosmological initial conditions and are replaced with a larger number of lower-mass particles (including gas if

[10] Technically, the statistical nature of the primordial fluctuations depends on how they were generated. The most commonly accepted theory is that of cosmic inflation (Guth, 1981), which predicts Gaussian fluctuations.

this is a hydrodynamical simulation). Lastly, the masses of the particles outside the high-resolution region are progressively increased and their number density correspondingly decreased in order to maintain a realistic tidal field on large scales with the minimum computational cost. This is the technique used to simulate the majority of the FABLE simulations, as detailed in Chapter 2 (Section 2.2.1).

1.8.1.4 *Halo finders*

Once the simulation is complete (or after the state of the simulation at a given time is output as a 'snapshot'), the final result is essentially just a box of numbers describing the properties of individual particles (e.g. their positions and velocities). At this stage it is often helpful to identify the gravitationally-bound haloes of particles corresponding to galaxies, clusters and other structures. The first step of this process is typically achieved using the friends-of-friends (FoF) algorithm, which defines haloes as structures whose particles are separated by a distance less than a specified linking length in units of the mean inter-particle distance (Davis et al., 1985).

The second step is usually to identify gravitationally self-bound "subhaloes" within the FoF haloes that correspond to, for example, galaxies within a galaxy group or cluster. A common technique for this purpose – and the one used in this thesis – is the SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009). Using an iterative gravitational unbinding procedure, SUBFIND identifies a single primary subhalo (also referred to as the "central" or "main" halo) per FoF group, which is by construction the most massive subhalo, and a number of secondary ("satellite") subhaloes. In the context of a galaxy cluster, the primary subhalo corresponds to the BCG and the secondary haloes to satellite galaxies.

1.8.2 *Semi-analytic modelling*

Results from the first dark matter simulations prompted the development of semi-analytic models that try to construct a coherent set of analytic approximations describing the physical processes relevant to galaxy formation (e.g. White & Frenk 1991; Cole 1991). Since then, numerous studies have extended and improved the original models (see Baugh 2006 for a review of semi-analytic techniques). The primary advantage of the semi-analytic approach compared with hydrodynamical simulations is that it is computationally much less expensive. This enables faster exploration of the parameter/model space and the construction of much larger samples of objects. The major disadvantage of the method is that it involves a particularly large degree of approximation. For example, whereas fully hydrody-

namic simulations follow the evolution of gas and dark matter self-consistently, semi-analytic models are forced to make restrictive assumptions about key dynamical processes affecting the gas, such as mergers and inflows. Comparison studies of semi-analytic versus hydrodynamic simulations find reasonable agreement in terms of simplified physics (e.g. hydrodynamics and cooling only; Benson et al. 2001; Yoshida et al. 2002; Helly et al. 2003), although recent advances in galaxy formation studies using cosmological hydrodynamical simulations suggests a number of discrepancies in their predicted galaxy properties that are generally attributed to improper assumptions and missing physics in the semi-analytic models (e.g. Hirschmann et al. 2012; Guo et al. 2016; Mitchell et al. 2018).

1.8.3 *Hydrodynamical simulations*

Modelling dark matter alone is insufficient to accurately describe the formation of galaxies. After all, what we observe as a galaxy is not dark matter but radiation originating from baryons. Unfortunately, adding baryonic material into an N-body simulation is not a straightforward task. Simulating the motions of gas (the dominant baryonic component) requires solving not only the equations of gravity but also the equations of hydrodynamics, which describe the motion of fluids. This involves, at the very least, computing the pressure forces acting on the gas and tracking its internal energy.

The two major approaches to modelling gas in hydrodynamical simulations are particle-based methods and mesh-based methods. In particle-based methods, the continuous gas distribution is discretised into a set of particles similar to those used to model the dark matter. Properties of the gas at a given point (e.g. its temperature) can then be interpolated from the particle distribution. The most common particle-based method is the Smoothed Particle Hydrodynamics scheme (SPH; see Springel 2010a for a review), which I describe in the next section. Mesh-based methods on the other hand discretise the simulation volume into (usually cubic) cells and calculate the flow of gas between neighbours. The most common class of mesh-based methods are Eulerian methods that employ a fixed, unmoving mesh as described below.

1.8.3.1 *Smoothed Particle Hydrodynamics (SPH)*

In SPH the properties of the fluid at a given point are calculated as a weighted average over nearby particles. The sum is weighted by i) the mass of the particles, and ii) their distance according to a specific function, known as a kernel. Typically the kernel is zero outside a given "smoothing length", which is usually chosen to

enclose a certain number of nearest neighbours. SPH therefore provides a spatially smoothed estimate of the fluid properties. The change in density, velocity and internal energy of each particle in the next time step can then be calculated from the spatial derivative of the smoothed field using the equations of hydrodynamics.

One of the major advantages of SPH is that particles will cluster in regions of higher density, yielding higher resolution where it is needed (e.g. in galaxies, in particular their dense central regions). In addition, particle-based methods like SPH couple easily to existing gravity solvers developed for N-body simulations. The major disadvantage of traditional SPH schemes is their inaccurate treatment of shock waves and contact discontinuities (e.g. Agertz et al. 2007; Sijacki et al. 2012). The latter, for example, leads to artificial suppression of fluid instabilities such as the Kelvin-Helmholtz instability, which are thought to be important mechanisms by which gas is stripped from galaxies moving through the ICM (e.g. Zavala et al. 2012). On the other hand, modern SPH schemes are able to mitigate these issues, for example by introducing artificial viscosity (Dolag et al., 2005) or thermal conductivity (Price, 2008).

1.8.3.2 *Eulerian methods*

In Eulerian methods the simplest mesh geometry is a grid of cubic cells of fixed size. More complicated geometries are also possible, although this is usually only an advantage for problems with a specific symmetry. A downside of using a grid of fixed resolution is that, unlike SPH, the resolution of the hydrodynamic treatment does not adapt to changes in density. To overcome this problem it is typical for simulations to use the "Adaptive Mesh Refinement" (AMR) technique to increase the resolution in dense regions by splitting cells that contain a mass above a certain threshold into multiple smaller cells. This can be done recursively to greatly increase the dynamic range of the simulation, albeit at the cost of extra computation time.

The primary advantage of mesh-based methods compared to SPH is a more accurate treatment of shocks and fluid instabilities. However, unlike SPH, Eulerian methods are not Galilean invariant, which means they can lose accuracy in the presence of bulk velocities (e.g. Tasker et al. 2008; Wadsley et al. 2008; Springel 2010b). This is of particular concern in galaxy formation simulations in which galaxies move at large speeds relative to each other and, for cluster galaxies, with respect to the ICM. In addition, AMR simulations of structure formation can underestimate the abundance of low-mass haloes at late times due to the discontinuous jumps in resolution implicit in the adaptive refinement method (O'Shea et al., 2005; Heitmann et al., 2008). On the other hand, modern AMR

codes are able to reach high enough resolution that these issues are significantly mitigated.

1.8.3.3 AREPO: *Moving mesh hydrodynamics*

Clearly, both SPH and Eulerian methods have a number of inherent advantages and disadvantages (see also Agertz et al. 2007 and Wadsley et al. 2008 for a detailed comparison between the two techniques). An alternative approach is the moving-mesh hydrodynamics scheme employed by the AREPO code, which combines many of the advantages of particle- and mesh-based methods whilst mitigating their respective disadvantages (Springel, 2010b).

In AREPO the equations of hydrodynamics are solved on an unstructured mesh defined by the Voronoi tessellation of a set of discrete points[11] that move with the velocity of the local flow. This method is largely Galilean invariant, which allows the proper treatment of the supersonic bulk flows that are common in cosmological simulations. Furthermore, the mesh-generating points cluster in dense regions similar to SPH, thereby automatically adapting the spatial resolution of the simulation to where it is needed. The moving-mesh scheme also inherits the high accuracy of Eulerian methods in the treatment of shocks, contact discontinuities and fluid instabilities, which are important for shock heating, gas stripping from satellites and the production of turbulence and entropy in the ICM (see Kereš et al. 2012; Sijacki et al. 2012; Vogelsberger et al. 2012; Nelson et al. 2013).

1.8.4 *Modelling galaxies and clusters in hydrodynamical simulations*

For hydrodynamical simulations to effectively model the complex process of galaxy and galaxy cluster formation and evolution they require a broad range of additional physics on top of gravity and hydrodynamics. This endeavour has been a major topic in astrophysics over the past three decades. The first three-dimensional simulations of dark matter and gas in the context of galaxy formation were described in Katz & Gunn (1991) and the first galaxy formation simulation run from cosmological, cold dark matter initial conditions was that of Navarro & White (1994). At the time, the typical resolution of such simulations was on the order of five particles per comoving $(1\,\mathrm{Mpc})^3$ volume. In the intervening decades, rapid increases in the power of modern supercomputers and advances in the efficiency and fidelity of numerical codes have facilitated hydrodynamic galaxy formation simulations of cosmological volumes with several thousand resolution

[11] The Voronoi tessellation of a set of points consists of non-overlapping cells centred on each point such that each cell contains the region of space that is closer to the given point than any other.

elements per comoving $(1\,\mathrm{Mpc})^3$ (e.g. Vogelsberger et al. 2014a; Schaye et al. 2015; Khandai et al. 2015), with some recent full-physics simulations reaching as high as $\sim 10^5$ Mpc^{-3} (IllustrisTNG; Nelson et al. 2019; Pillepich et al. 2019).

Even with this rapid increase in resolution, the scales on which most astrophysical processes operate are still well below the resolution of current cosmological simulations. As such, galaxy formation physics are often modelled in a "sub-grid" fashion, wherein the effects of physical processes on unresolved scales are approximated using a relatively simple model based on theoretical expectations and/or empirical evidence.

Some of these effects are theoretically well-understood, such as radiative cooling and photoionization. The addition of primordial and metal-line cooling in hydrodynamical simulations is relatively straightforward and can be achieved by tabulating the rate at which gas of a given density, temperature and metallicity is expected to radiate energy and interpolating from these tables as necessary. Modern codes also take into account the suppression of the cooling rate and increased heating due to photoionization of the gas by the UV background, which is typically assumed to be time-varying but spatially uniform (e.g. Wiersma et al. 2009a).

On the other hand, the effects of many other physical processes are much more difficult to predict a priori. As such, the details of the sub-grid modelling must be estimated empirically or tuned to match observational results. For example, star formation models must assume a form for the IMF typically estimated from the observed dynamics of stars in the Milky Way or nearby galaxies (e.g. Cappellari et al. 2006). This choice has a non-negligible impact on the recycling of metal-rich material from ageing stellar populations, which plays a central role in the evolution of a galaxy's gas content. As a further example, the sub-grid model parameters that control the driving rate of supernovae-driven galactic winds in simulations significantly influences their ability to regulate star formation (e.g. Vogelsberger et al. 2013). These parameters cannot be easily predicted and hence their values are typically calibrated such that the simulation matches a chosen observable, for example, the abundance of low-mass galaxies.

The calibration of sub-grid parameters does not guarantee good agreement with observations however. For example, simulations have shown that supernovae feedback alone is insufficient to regulate the growth of massive galaxies (e.g. Crain et al. 2009; McCarthy et al. 2012; Kannan et al. 2014). Simulations have also demonstrated a potential solution to this problem however in the form of feedback from actively accreting black holes (AGN).

1.8.4.1 *Black holes and AGN feedback*

Black hole growth and AGN feedback in hydrodynamical simulations were first modelled by Di Matteo et al. (2005) and Springel et al. (2005a) in simulations of galaxy mergers. In their implementation, black holes are represented by collisionless particles that grow in mass by accreting gas from their environment at a rate specified by the Bondi-Hoyle-Lyttleton parametrisation, which describes the rate at which a point mass moving through a uniform gas cloud accretes mass (see Edgar 2004 for a useful summary). It is assumed that a constant fraction of the accreted mass is converted into radiation by the black hole. This fraction, called the radiative efficiency, is typically chosen to be ~ 10–20 per cent based on theoretical expectations. The maximum accretion rate is limited to the Eddington rate, which is the accretion rate at which the pressure of the emitted radiation balances the inward gravitational force. In this model AGN feedback consists of continuous and isotropic injection of thermal energy into the surrounding gas at a fraction, ϵ_f, of the radiative luminosity of the black hole. Di Matteo et al. (2005) found that this scheme, with a feedback efficiency of $\epsilon_f \sim 0.05$, is able to reproduce the observed relation between black hole mass and stellar velocity dispersion in galaxies. Contemporary models for black growth and AGN feedback typically use some variation on this approach. Although, as for supernovae feedback, the parameters of the model must be calibrated to match some observable, for example the abundance of massive galaxies (e.g. Puchwein & Springel 2013; Vogelsberger et al. 2013; Crain et al. 2015; Steinborn et al. 2015; McCarthy et al. 2017).

1.8.4.2 *Recent progress*

AGN feedback in particular has proven to be a key ingredient in hydrodynamical galaxy and galaxy cluster simulations (e.g. Croton et al. 2006; Sijacki et al. 2007; Booth & Schaye 2009; McCarthy et al. 2010, 2011), in addition to important physical processes such as radiative cooling, star formation, stellar evolution, chemical enrichment and stellar feedback among others (see Borgani & Kravtsov 2011 for a review). The combination of this diverse ensemble of physical processes, evolved self-consistently alongside gravity and hydrodynamics, represents one of the major advantages of galaxy formation simulations over other techniques such as semi-analytic modelling. Yet this complex interplay of non-linear processes, acting over an enormous range of scales (from many Mpc to sub-pc scales), also represents a formidable computational challenge. Indeed, performing such simulations in a truly self-consistent manner remains a distant goal. Nevertheless, remarkable progress has already been made in this direction in recent years thanks to increases in computing power and a huge effort in the development of sub-grid

modelling.

So far, efforts in galaxy formation simulations have typically progressed along one of two parallel paths: high-resolution simulations on the scale of a single galaxy or smaller focused on specific phenomena (e.g. Guedes et al. 2011; Hopkins et al. 2014; Marinacci et al. 2014; Wang et al. 2015; Agertz & Kravtsov 2016; Christensen et al. 2016; Grand et al. 2017); or large cosmological volumes that attempt to reproduce the integral properties of entire galaxy populations arising from cosmological initial conditions (e.g. Dubois et al. 2014; Hirschmann et al. 2014; Vogelsberger et al. 2014a; Khandai et al. 2015; Tremmel et al. 2017; Pillepich et al. 2018a). The latter are limited in spatial and mass resolution to scales on the order of a kiloparsec and masses of $10^6\,M_\odot$. Nevertheless, in recent years projects such as EAGLE (Schaye et al., 2015; Crain et al., 2015) and Illustris (Vogelsberger et al., 2014b; Genel et al., 2014) have demonstrated that galaxy formation simulations tailored to such resolutions can reproduce many of the structural properties and scaling relations of observed galaxies across a wide range of masses.

On the other hand, the typical volume of these simulations ($\lesssim 100$ Mpc on a side) limits their usefulness for the study of rare objects such as galaxy clusters. Fortunately, in parallel to the advancements in the modelling of galaxies in hydrodynamical simulations, a number of studies have also focused their attention on the modelling of galaxy clusters and groups. Similar to the approach described above, these efforts have typically proceeded along one of two tracks: relatively low resolution simulations of large, statistical samples of group- and cluster-scale haloes (e.g. Sembolini et al. 2013; Le Brun et al. 2014; Dolag et al. 2016; McCarthy et al. 2017; Barnes et al. 2017a; Cui et al. 2018); or simulations of small samples of clusters and/or groups focused on specific problems, for example, the numerical implementation (e.g. Frenk et al. 1999; Dolag et al. 2005; Voit et al. 2005; Mitchell et al. 2009; Sembolini et al. 2016a,b), the impact of AGN feedback on the thermodynamic properties of the ICM (e.g. Sijacki et al. 2007; Puchwein et al. 2008; Fabjan et al. 2010; McCarthy et al. 2010, 2011; Pike et al. 2014; Planelles et al. 2014) or the growth of the stellar mass content (e.g. Puchwein et al. 2010; Martizzi et al. 2012, 2014; Remus et al. 2017; Pillepich et al. 2018b).

These simulations have met with varying degrees of success. The majority of studies have found that AGN feedback causes a significant reduction in the gas mass fractions of galaxy groups and low-mass clusters, in good agreement with observations (e.g. Sijacki et al. 2007; Puchwein et al. 2008; McCarthy et al. 2010, 2011; but see also Battaglia et al. 2013; Planelles et al. 2013; Hahn et al. 2017). In turn this has led to a considerable improvement in the modelling of scaling relations between various X-ray and SZ properties of clusters (e.g. Puchwein et al.

2008; Le Brun et al. 2014; Pike et al. 2014; Planelles et al. 2014; McCarthy et al. 2017; Barnes et al. 2017a; Truong et al. 2018). However, whilst this means that reasonable agreement with observations can be achieved in terms of the *global* properties of simulated clusters, it does not guarantee that the ICM is realistically distributed. Indeed, reproducing the observed thermodynamic profiles of gas in clusters (particularly their central regions) has proven a significant challenge, with many simulations struggling to achieve this goal (e.g. Mitchell et al. 2009; Borgani & Kravtsov 2011; Pike et al. 2014; Planelles et al. 2014; Barnes et al. 2017b). A handful of recent studies have made progress in this regard by overcoming a long-standing problem in the modelling of cluster thermodynamic profiles by producing distinct populations of so-called 'cool-core' and 'non-cool-core clusters' with ICM profiles in relatively good agreement with observations (e.g. Rasia et al. 2015; Hahn et al. 2017; Barnes et al. 2018b; Planelles et al. 2017; Barnes et al. 2018a), albeit with some debate over the mechanism responsible for the production and maintenance of cool cores (see e.g. discussion in Barnes et al. 2018b).

In terms of the stellar mass content of groups and clusters, the results of cluster simulations are consistent with simulations of field galaxy populations in the sense that some form of AGN feedback is typically required to produce stellar mass fractions in agreement with observations (e.g. McCarthy et al. 2010; Puchwein et al. 2010; Planelles et al. 2013; Pillepich et al. 2018a). Combined with a prescription for stellar feedback, some recent simulations have even reproduced the stellar mass function of satellite galaxies in clusters (e.g. Bahé et al. 2017; Cui et al. 2018). Along similar lines, AGN feedback is required to produce BCGs with realistic stellar masses (e.g. Martizzi et al. 2012, 2014; Ragone-Figueroa et al. 2018), although even with AGN feedback many simulations still produce overly massive BCGs (e.g. Puchwein et al. 2010; Ragone-Figueroa et al. 2013; Bahé et al. 2017; Cui et al. 2018; Pillepich et al. 2018b).

Projects such as MACSIS (Barnes et al. 2017a; an extension of the BAHAMAS project; McCarthy et al. 2017) and The Three Hundred project (Cui et al., 2018) have produced large catalogues of simulated galaxy clusters as useful tests of large-scale structure formation and cosmology. The simulations are a good match to a range of observations, especially the X-ray and SZ scaling relations that are of particular relevance to cluster cosmology. However there is still significant room for improvement, in particular the evolution of star formation rates in MACSIS/BAHAMAS (McCarthy et al., 2017) and the total stellar mass of galaxy groups for The Three Hundred project (Cui et al., 2018). Furthermore, the relatively low mass resolution of these simulations allows the study of only relative massive galaxies with stellar masses $M_\star \gtrsim 10^{10} M_\odot$.

Recently, a handful of hydrodynamical simulations have attempted to combine recent advancements in galaxy and galaxy cluster formation simulations in an effort to reproduce *simultaneously* the observed properties of galaxies, groups and clusters across the full mass range. Notable examples include C-EAGLE (Barnes et al., 2017b; Bahé et al., 2017), which employs the same resolution and galaxy formation model as EAGLE to simulate 30 galaxy clusters across a wide mass range, and the IllustrisTNG project (Marinacci et al., 2018; Naiman et al., 2018; Nelson et al., 2018; Pillepich et al., 2018b; Springel et al., 2018), which updates the Illustris galaxy formation model to simulate three large, uniformly-sampled cosmological volumes containing more than two hundred galaxy clusters.

The FABLE project represents a complementary step forward in this direction. Working independently of IllustrisTNG, FABLE employs an updated and recalibrated version of the Illustris galaxy formation model that addresses many of the previous shortcomings of Illustris on group and cluster scales, whilst maintaining good agreement with observations on galaxy scales. The FABLE suite of simulations includes a large, uniformly-sampled cosmological volume containing several thousand galaxies and a series of simulations of individual galaxy groups and clusters and their environments, totalling roughly twenty galaxy clusters ($M_{500} \gtrsim 10^{14} M_{\odot}$) and a similar number of galaxy groups ($3 \times 10^{13} M_{\odot} \lesssim M_{500} \lesssim 10^{14} M_{\odot}$). In this thesis I will present an investigation into the diverse properties of FABLE systems with comparison to observations and predictions from other simulation studies. I will use these comparisons to provide possible interpretations for observed phenomena, to make predictions for future observations, and to highlight areas for improvement in the physical modelling of galaxy clusters and groups within cosmological hydrodynamical simulations.

1.9 THESIS SETUP

The focus of this thesis is the study of galaxy clusters and groups in the FABLE simulations, in particular the observable properties of the hot ICM and their co-dependence as captured by the X-ray and SZ scaling relations, as well as the stellar content of clusters and their constituent galaxies. Below I summarise the content of each chapter.

- In Chapter 2 I describe the FABLE project in detail, including its motivation, development and the properties of the simulations. I perform a detailed comparison to observational constraints of the galaxy stellar mass function and the baryon fractions and X-ray and SZ scaling relations of groups and

clusters at low-redshift, taking care to make as close to a like-for-like comparison as possible. I demonstrate good agreement of the simulations with observations in a number of areas, in particular the stellar mass content of clusters, groups and galaxies and the scaling relations between X-ray luminosity or SZ signal and total mass. I discuss possible explanations for the remaining discrepancies, in particular a potential X-ray hydrostatic mass bias in observed samples.

- In Chapter 3 I investigate the redshift evolution of the X-ray and SZ scaling relations in FABLE in comparison to predictions from other recent simulations. For all of the examined relations FABLE predicts significant deviations from the simple self-similar expectations in terms of their slope and the redshift evolution of their normalisation. This has important consequences for the use of mass proxies and cluster abundance measurements in constraining cosmology with upcoming samples of high-redshift clusters, which I demonstrate by assessing the sensitivity of SZ cluster counts to the underlying SZ signal–total mass relation.

- In Chapter 4 I study the mass and redshift dependence of the baryon content in FABLE clusters and groups, finding little evolution in their total gas or stellar mass at $z \lesssim 1$, consistent with recent observations. I also investigate the population of BCGs in FABLE, including their total stellar mass and stellar mass profiles and how this varies with the cluster mass and/or redshift.

- In Chapter 5 I summarise the main conclusions of this work and discuss further avenues for research.

2 | THE FABLE SIMULATIONS

In this chapter I introduce the *Feedback Acting on Baryons in Large-scale Environments* (FABLE) suite of cosmological hydrodynamical simulations of galaxies, groups and clusters. The simulations use the AREPO moving-mesh code with a set of physical models for galaxy formation based on the successful Illustris simulation, but with updated AGN and supernovae feedback models. This model simultaneously reproduces the observed redshift evolution of the galaxy stellar mass function together with the stellar and gas mass fractions of local groups and clusters across a wide range of halo masses. Focusing on the $z \sim 0$ properties of groups and clusters, I demonstrate very good agreement with a range of observed scaling relations, including the X-ray luminosity–total mass and gas mass relations as well as the total mass–temperature and Sunyaev-Zel'dovich flux–mass relations. Careful comparison with scaling relations based on X-ray hydrostatic masses as opposed to weak lensing-derived masses reveals some discrepancies, which hint towards a non-negligible X-ray mass bias in observed samples. I also show that radial profiles of density, pressure and temperature of the simulated intracluster medium are in very good agreement with observations, in particular for $r > 0.3\, r_{500}$. In the innermost regions however the simulated clusters possess too large entropy cores, which indicates that a more sophisticated modelling of the physics of AGN feedback may be required to accurately reproduce the observed populations of cool-core and non-cool-core clusters.

The work presented in this chapter has been published in collaboration with Ewald Puchwein, Sijing Shen and Debora Sijacki as:[1]

Henden N. A., Puchwein E., Shen S., Sijacki D., *The FABLE simulations: a feedback model for galaxies, groups, and clusters*, 2018, MNRAS, 479, 5385.

[1] I personally ran all of the simulations in the FABLE suite and performed the entirety of the analyses presented in this chapter. The simulations were run using the AREPO code provided by Volker Springel and collaborators and initial conditions were produced by my supervisors, Ewald Puchwein and Debora Sijacki. The simulations used in the calibration of the FABLE model are based on simulations performed by Debora Sijacki and Sijing Shen.

2.1 BACKGROUND

Ongoing and forthcoming surveys are set to greatly extend the number of known galaxy clusters and groups. In X-rays, the upcoming eROSITA telescope (Merloni et al., 2012; Pillepich et al., 2012) is expected to detect as many as one hundred thousand groups and clusters of galaxies out to $z \sim 1$ while observations of the Sunyaev–Zel'dovich effect (SZ) with SPT-3G (Benson et al., 2014) and ACTpol (Henderson et al., 2016) extend the search to higher redshift. Detailed follow-up observations will be possible at X-ray wavelengths with missions such as *XARM* (replacement for *Hitomi*; Takahashi et al. 2010, 2016) and *Athena* (Nandra et al., 2013) and at optical and near-infrared wavelengths with ground-based surveys such as the Dark Energy Survey (Dark Energy Survey Collaboration, 2005) and the Large Synoptic Survey Telescope (LSST Dark Energy Science Collaboration, 2012), as well as future space-based missions such as Euclid (Laureijs et al., 2011).

This wealth of data holds great potential to provide stringent constraints on cluster physics and cosmological parameters, such as the amplitude and slope of the matter power spectrum and the densities of baryons, dark matter and dark energy (see Allen et al. 2011; Kravtsov & Borgani 2012; Planelles et al. 2015 for recent reviews). However, the cosmological interpretation of observed cluster data is crucially dependent on the degree to which we can connect cluster observables to theoretical predictions for a given cosmological model. Simulations of cosmic structure formation play an important role in this process. At the most basic level, they provide predictions for the abundance of clusters as a function of their mass for a particular cosmology (e.g. Cohn & White, 2008; Tinker et al., 2008; Courtin et al., 2010; Crocce et al., 2010; Bhattacharya et al., 2011; Murray et al., 2013b; Watson et al., 2013). Yet as larger cluster surveys reduce the statistical uncertainty, we are increasingly limited by our incomplete understanding of cluster physics and its impact on the relation between cluster observables (such as X-ray luminosity, temperature or SZ flux) and mass (e.g. Rudd et al., 2008; Semboloni et al., 2011; van Daalen et al., 2011; Cui et al., 2014b; Cusworth et al., 2014; Velliscig et al., 2014; Henson et al., 2017; Chisari et al., 2018).

Cosmological hydrodynamical simulations can aid in understanding and constraining these relations, as well as explaining potential systematic biases in observations. As an example, the most common method for measuring cluster masses is by analysing X-ray data under the assumption that the X-ray emitting gas is in hydrostatic equilibrium with the gravitational potential of the cluster. This method has been reported to systematically underestimate the true halo mass (e.g. Miralda-Escude & Babul 1995; Wu & Fang 1997; Mahdavi et al. 2008;

Richard et al. 2010; Mahdavi et al. 2013; Foëx et al. 2017), however, there is still no consensus on the exact magnitude of the effect or indeed whether such a bias exists at all (e.g. Zhang et al., 2010; Gruen et al., 2014; Applegate et al., 2016; Planck Collaboration XXIV, 2016; Rines et al., 2016). Furthermore, as cluster surveys expand in size and depth, it will be vital to understand potential selection biases associated with a given observable. For example, X-ray surveys may be biased toward cool-core clusters with strong central X-ray emission, the prevalence and evolution of which remains uncertain (e.g. Andrade-Santos et al., 2017; Rossetti et al., 2017). Cosmological hydrodynamical simulations are in a unique position to quantify such potential biases and therefore facilitate the use of clusters as precise cosmological probes.

The hydrodynamical modelling of cluster formation has progressed considerably in recent years, largely due to increases in computing power and improvements in sub-grid modelling (see Section 1.8.4). This has facilitated hydrodynamical simulations aimed at reproducing various observational properties of clusters, such as the density, temperature and metallicity profiles of the intracluster medium (e.g. Planelles et al. 2014; Barnes et al. 2017a,b; Biffi et al. 2017; Vogelsberger et al. 2018), the cool-core/non-cool-core dichotomy (e.g. Rasia et al. 2015; Hahn et al. 2017; Barnes et al. 2018b,a) and the X-ray and SZ scaling relations (e.g. Pike et al. 2014; Planelles et al. 2014), including their redshift evolution (e.g. Fabjan et al. 2011; Le Brun et al. 2016; Truong et al. 2018). A small number of groups have also produced hydrodynamical simulations of statistical samples of massive haloes useful for cluster cosmology (Le Brun et al., 2014; Dolag et al., 2016; Barnes et al., 2017a; McCarthy et al., 2017). One of the limitations of these works however is that they lack the numerical resolution needed to resolve the detailed structure of the cluster galaxies. Given that galaxies are often used as observational tracers of large-scale structure, it is important to work towards reproducing the properties of the galaxy populations of clusters in addition to their global stellar, gas and halo properties.

As progress has been made in the modelling of cluster formation, so too has there been a marked increase in the realism of simulated field galaxy populations. A number of groups have produced high-resolution cosmological hydrodynamical simulations of galaxy formation capable of reproducing a range of key observations of galaxies, such as their sizes, morphologies, passive fractions and the build-up of their stellar mass. Notable examples include Illustris (Vogelsberger et al., 2014a), EAGLE (Schaye et al., 2015), Horizon-AGN (Dubois et al., 2014) and MassiveBlack-II (Khandai et al., 2015). The computational expense associated with the high resolution and complex sub-grid models of these simulations limits their

total volume. As a result, few, if any, galaxy clusters exist within the simulation volume and it is difficult to assess the ability of the galaxy formation model to reproduce observations of massive collapsed structures.

Recently, the C-EAGLE (Barnes et al., 2017b; Bahé et al., 2017) and IllustrisTNG (Pillepich et al., 2018b) projects have combined these two approaches and produced high-resolution simulations that resolve the full dynamic range from galaxies to massive clusters. The C-EAGLE project consists of "zoom-in" simulations of 30 galaxy clusters in the mass range $M_{200} = 10^{14} - 10^{15.4} M_{\odot}$ simulated with the same galaxy formation model as the EAGLE simulation and at the same resolution. The C-EAGLE clusters are a good match to a number of cluster observables, including the satellite stellar mass function, the total stellar and metal content and the scaling of X-ray spectroscopic temperature and SZ flux with total mass. Conversely, the clusters are too gas rich and possess too high central temperatures. Their results imply that improved agreement with observations will require revision of their model for AGN feedback (see discussion in Barnes et al. 2017b). The IllustrisTNG project consists of three large, uniformly-sampled cosmological volumes of approximately 50, 100 and 300 Mpc on a side simulated with magneto-hydrodynamics (Pillepich et al., 2018a). The largest volume contains numerous low-mass clusters, with a few objects reaching masses $M_{200} \sim 10^{15} M_{\odot}$ at $z = 0$. IllustrisTNG is a follow-up project of the Illustris simulation (Vogelsberger et al., 2013, 2014a; Genel et al., 2014) and improves upon Illustris by extending the explorable mass range of simulated haloes and revising the astrophysical modelling to address some of the shortcomings of Illustris. The Illustris simulation was successful in reproducing a broad range of observations of galaxy populations at various redshifts, including the observed range of galaxy morphologies and the evolution of galaxy specific star formation rates. However, various tensions with observations remained, especially on the scale of galaxy groups and (low-mass) clusters. In particular, the AGN feedback in Illustris ejected too much gas from low-redshift, massive haloes ($M_{500} \approx 10^{13} - 10^{14} M_{\odot}$; Genel et al. 2014). IllustrisTNG introduce a new model for AGN feedback which is able to suppress star formation in massive haloes without removing too much gas (Weinberger et al., 2017a). Combined with changes to other aspects of the modelling, such as modified galactic winds and the introduction of magnetic fields (Pillepich et al., 2018a), this has allowed IllustrisTNG to reproduce a range of observables properties of galaxies, groups and clusters.

The *Feedback Acting on Baryons in Large-scale Environments* (FABLE) project presented here shares a similar motivation to IllustrisTNG. We have worked independently to build a suite of simulations based upon the framework of the successful

Illustris project but improve upon the agreement with observations on scales larger than galaxies, e.g., with constraints on the gas content of massive haloes. By using zoom-in simulations we also extend the comparison with observations to massive clusters, which were not present in the original Illustris simulation. Like IllustrisTNG we have updated the Illustris models for galactic winds and AGN feedback. We find that our changes improve the realism of groups and clusters significantly. The FABLE simulations consist of a uniformly-sampled cosmological volume about 60 Mpc on a side and a series of zoom-in simulations of groups and clusters approximately uniformly spaced in logarithmic halo mass. In this chapter we focus on the $z = 0$ properties of FABLE groups and clusters in comparison to observations and demonstrate good agreement in a number of key areas.

The outline of this chapter is as follows. Section 2.2 describes the characteristics of the FABLE simulations and Section 2.3 describes our methods for comparing the simulations with observations. We then present the galaxy stellar mass function at different redshifts (Section 2.4), the global properties of massive haloes (Section 2.5) and the hot gas profiles of groups and clusters (Section 2.6). We discuss our results in Section 2.7 and summarise our findings in Section 2.8.

2.2 SIMULATIONS

2.2.1 *Basic simulation properties*

We utilise the cosmological hydrodynamic moving-mesh code AREPO (Springel, 2010b), which solves the Euler equations of hydrodynamics on a quasi-Lagrangian moving Voronoi mesh. On top of gravity, hydrodynamics and a spatially uniform ionizing background, the FABLE simulations employ a set of sub-grid models for processes important for galaxy formation. The majority of our sub-grid models are unchanged from Illustris and are described in full detail in Vogelsberger et al. (2013) and Torrey et al. (2014). These include models for star formation (Springel & Hernquist, 2003a; Springel et al., 2005a), radiative cooling (Katz et al., 1996; Wiersma et al., 2009a) and chemical enrichment (Wiersma et al., 2009b). The sub-grid models for feedback from stars (Vogelsberger et al., 2013) and AGN (Di Matteo et al., 2005; Springel et al., 2005a; Sijacki et al., 2007) have been modified from the Illustris models and are described in Sections 2.2.3 and 2.2.4. The parameters of the feedback models have been calibrated to reproduce observations of the local galaxy stellar mass function (Section 2.4) and the gas mass fractions of massive haloes (Section 2.5). This same calibration strategy was adopted for the BAHAMAS simulations (McCarthy et al., 2017), which successfully reproduce

a broad range of observed hot gas and stellar properties of massive systems. The calibration process involved a series of periodic boxes of length 40 h^{-1} Mpc simulated with different parametrizations of stellar and AGN feedback. We note that, as this volume is relatively small, these simulations did not contain massive clusters and therefore gas fractions were initially matched to observations on group-scales ($M_{500} \lesssim 10^{14} M_{\odot}$). Results from a sample of these calibration simulations are presented in Appendix A. The FABLE suite of simulations consists of the calibration volume and a series of zoom-in simulations of individual galaxy groups and clusters, which apply our calibrated model to higher mass objects.

The 40 h^{-1} (comoving) Mpc periodic box was evolved to $z = 0$ from initial conditions based on a Planck cosmology (Planck Collaboration XIII, 2016) with cosmological parameters $\Omega_{\Lambda} = 0.6935$, $\Omega_{\rm M} = 0.3065$, $\Omega_{\rm b} = 0.0483$, $\sigma_8 = 0.8154$, $n_s = 0.9681$ and $H_0 = 67.9$ km s^{-1} Mpc^{-1} $= h\times$100 km s^{-1} Mpc^{-1}. The simulation follows 512^3 dark matter particles of mass $m_{\rm DM} = 3.4\times10^7\ h^{-1}$ M$_{\odot}$ and approximately 512^3 baryonic resolution elements (gas cells and star/BH particles) of typical mass $\overline{m}_{\rm b} = 6.4\times10^6\ h^{-1}$ M$_{\odot}$. The gravitational softening length was fixed to $2.393\ h^{-1}$ kpc in physical coordinates below $z = 5$ and fixed in comoving coordinates at higher redshifts. This choice is consistent with the $z = 0$ optimal softening length for low-mass cluster galaxies according to the empirical rule determined by Power et al. (2003)[2].

For our series of zoom-in simulations, individual groups and clusters were chosen from a collisionless (dark matter-only) parent simulation and re-simulated at high resolution. The zoom-in regions were drawn from Millennium XXL, a periodic box of side length $3\ h^{-1}$ Gpc (Angulo et al., 2012). Six systems were selected logarithmically-spaced in mass spanning the mass range from groups ($\sim 10^{13} M_{\odot}$) to massive clusters ($\sim 10^{15} M_{\odot}$). These systems were selected only by their total mass within the parent simulation, with no prior knowledge regarding, for example, their dynamical state. The high-resolution regions were chosen such that they are free from lower resolution particles out to approximately $5\, r_{500}$ at $z = 0$. Dark matter particles in the high-resolution region have a mass of $m_{\rm DM} = 5.5\times10^7\ h^{-1}$ M$_{\odot}$. The gravitational softening length was fixed at $2.8125\ h^{-1}$ kpc in physical coordinates for $z \leq 5$ and fixed in comoving coordinates for $z > 5$. The softening lengths of boundary particles outside the high-resolution region were allowed to vary with their mass. Mode amplitudes in the initial conditions were scaled to a Planck cosmology (Planck Collaboration XIII, 2016) with $\Omega_{\Lambda} = 0.6911$, $\Omega_{\rm M} = 0.3089$, $\Omega_{\rm b} = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$ and $H_0 = 67.74$ km s^{-1} Mpc^{-1}.

[2]Note that this softening length is somewhat larger than used in Illustris.

In presenting our results we rescale the appropriate quantities to the cosmology of the periodic box described above, although we note that this is a very small effect due to the similarity between the cosmological parameters.

The simulations were processed on-the-fly with the friends-of-friends (FoF) and SUBFIND algorithms (Davis et al., 1985; Springel et al., 2001; Dolag et al., 2009) to identify gravitationally bound groups of dark matter, stars and gas ("haloes"), using a linking length of 0.2 times the mean dark matter inter-particle separation, and decompose them into self-bound substructures ("subhaloes").

2.2.2 *Cluster visualization*

In Fig. 2.1 we present a visualization of the gas properties of the most massive FABLE cluster at $z = 0$. The main panel shows the gas distribution in a $20 \times 20 \times 5$ Mpc slice centred on the gravitational potential minimum of the cluster. The gas surface density and temperature are encoded by the image brightness and hue/saturation, respectively, according to the colour map shown in the bottom-left of the figure. The structure of the cluster gas is clearly laid out, revealing an array of interconnected filaments and infalling galaxies and groups, examples of which are shown in the panels on the right hand side of the figure. During the formation of this cluster, a combination of adiabatic compression and shocks has heated the intracluster gas to high, X-ray emitting temperatures ($\gtrsim 10^7$ K). This is demonstrated in the middle right-hand panel, which shows the X-ray surface brightness within a region $2\,r_{500}$ on a side centred on the cluster.

2.2.3 *Star formation, stellar feedback and chemistry*

Star formation and chemistry are treated with the same implementation as in Illustris, as described in detail in Vogelsberger et al. (2013) and Torrey et al. (2014). In brief, the dense star-forming interstellar medium (ISM) is treated in a sub-resolution fashion using a slightly modified version of the Springel & Hernquist (2003a) sub-grid model. The ISM is modelled using an effective equation of state, where stars form stochastically from gas above a density threshold with a given star formation time scale. Stellar mass loss and metal enrichment are treated by calculating the mass and chemical composition of ejected material for each active star particle at each time step and returning it to the nearby gas.

Stellar feedback in the form of galactic outflows is modelled via wind particles launched stochastically from star forming gas with a velocity based on the local dark matter velocity dispersion (Vogelsberger et al., 2013). The wind particles

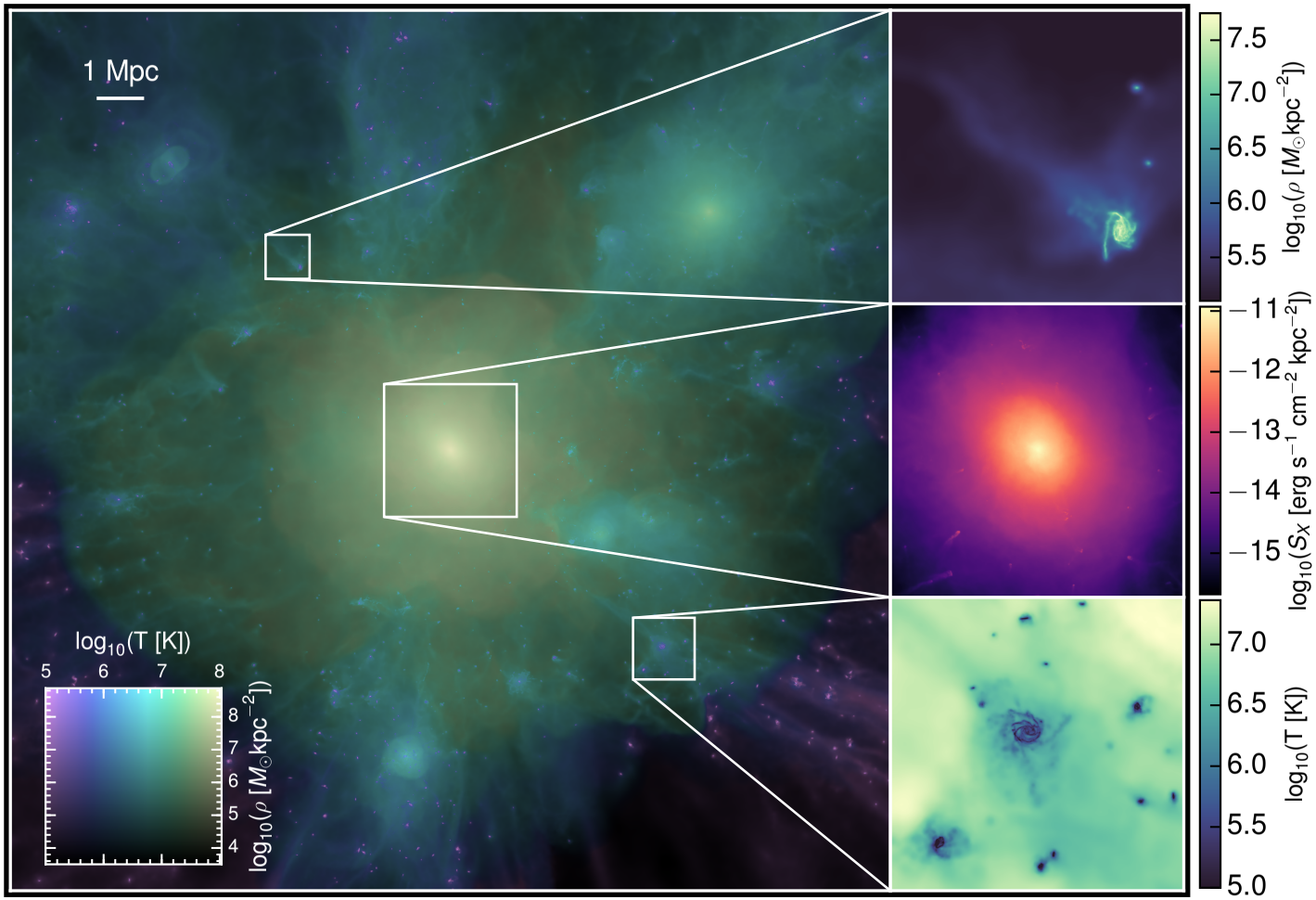


Figure 2.1: A visualization of the gas properties of a massive FABLE cluster with $M_{500} = 9 \times 10^{14} M_\odot$ at $z = 0$. The main panel shows a $20 \times 20 \times 5$ Mpc slice through the cluster with panels on the right hand side zooming in on individual galaxies and the cluster core. In the main panel the gas surface density is represented by the image brightness while the gas temperature corresponds to the hue/saturation of the image according to the inset colour scale. The bottom-right panel shows the gas temperature distribution of a group of galaxies of various sizes and morphologies. The top-right panel shows the surface density of gas surrounding an individual spiral galaxy on the cluster outskirts, revealing the onset of ram pressure stripping of cold galactic gas as the galaxy approaches the dense cluster environment. Lastly, in the middle-right panel we show the X-ray surface brightness in the 0.5–7 keV energy band within a region $2\ r_{500}$ on a side centred on the cluster.

are briefly decoupled from hydrodynamic interactions until they have left the local ISM and deposit their mass, metals, momentum and thermal energy into the surrounding gas. In Illustris, the energy given to the wind particles is purely kinetic. These cold winds push gas out of the galactic disc, preventing overcooling within the dense ISM and lowering the galactic star formation rate. However, in this model it is left fully up to hydrodynamical interactions to dissipate the kinetic energy to heat, which is often insufficient to prevent the ejected gas from quickly condensing back onto the galactic disc. We reduce this possibility by imparting one-third of the wind energy as thermal rather than kinetic. This slows the cooling of the ejected gas, causing it to remain outside of the dense ISM for a longer period of time, increasing the overall effectiveness of stellar feedback. The same method has been used in Marinacci et al. (2014) (50 per cent thermal energy), the Auriga galaxy simulations (50 per cent; Grand et al. 2017) and IllustrisTNG (10 per cent; Pillepich et al. 2018a).

2.2.4 *Black hole physics*

Black hole formation is modelled assuming that "seed" black holes (BHs) of mass $10^5 h^{-1} M_\odot$ form regularly enough that every halo above a mass threshold of $5 \times 10^{10} h^{-1} M_\odot$ contains a seed BH at its centre. Such haloes are identified by running a fast FoF algorithm on-the-fly. If a halo exceeds this mass threshold and does not already host a BH, the highest density gas cell in the halo is converted into a BH particle.

BHs are modelled as collisionless sink particles that are allowed to grow via mergers with other BHs and by accretion of ambient gas. The prescription for BH accretion is described in detail in Vogelsberger et al. (2013). Briefly, BHs accrete at an Eddington-limited Bondi-Hoyle-Lyttleton rate boosted by a constant factor $\alpha = 100$ (Hoyle & Lyttleton, 1939; Bondi & Hoyle, 1944; Springel et al., 2005a). A fraction $(1 - \epsilon_r)$ of the accreted mass is added to the mass of the BH, where $\epsilon_r = 0.1$ is the radiative efficiency. The remaining rest mass energy is made available as feedback energy.

For feedback from BHs we use an adapted form of the Illustris model. This model distinguishes between two states: a high accretion rate quasar-mode (Di Matteo et al., 2005; Springel et al., 2005a) and a low accretion rate radio-mode (Sijacki et al., 2007). The quasar-mode is dominant at high redshift while the radio-mode becomes increasingly important at lower redshifts as the average accretion rate diminishes.

In the quasar-mode of feedback, a fraction $\epsilon_f = 0.1$ of the available feedback

energy is coupled thermally and isotropically to the surrounding gas (see Springel et al. 2005a for details). In Illustris, BHs in the quasar-mode inject thermal energy into the surrounding gas *continuously*. However, this approach can result in artificial overcooling whereby too little thermal energy is injected into too much mass and the energy is quickly radiated away. For this reason, Booth & Schaye (2009) developed a model for thermal AGN feedback in smoothed particle hydrodynamics (SPH) simulations in which BHs store feedback energy until it is sufficient to raise the temperature of a given number of SPH particles by a given amount. This model was subsequently employed in the cosmological hydrodynamical simulations EAGLE (Schaye et al., 2015), cosmo-OWLS (Le Brun et al., 2014) and BAHAMAS (McCarthy et al., 2017). We take a similar approach to reduce numerical overcooling in our simulations by introducing a duty cycle for the quasar-mode in which feedback energy is accumulated over a given time period, $\Delta t = 25$ Myr, before being released in a single event.

At low accretion rates below a fraction $\chi_{\rm radio} = 0.01$ of the Eddington rate, we employ radio-mode feedback following Sijacki et al. (2007). In the radio-mode, hot buoyantly-rising bubbles are injected into the surrounding gas. The duty cycle of the radio-mode is controlled by the mass growth of the BH such that a bubble is injected once the BH has increased in mass by a fraction $\delta_{\rm BH} \equiv \delta M_{\rm BH}/M_{\rm BH}$. The energy content of the bubble is then $E_{\rm bub} = \epsilon_m \epsilon_r c^2 \delta M_{\rm BH}$ for which we choose a radio-mode coupling efficiency $\epsilon_m = 0.8$. The total radio-mode feedback efficiency as given by the product $\epsilon_m \epsilon_r$ is 8 per cent, similar to Illustris (7 per cent). For the duty cycle we set the threshold for bubble triggering to $\delta_{\rm BH} = 0.01$, much smaller than $\delta_{\rm BH} = 0.15$ as used in Illustris. Since the bubble energy is proportional to the increase in mass, this corresponds to more frequent but less energetic bubbles. For the bubble size and distance we assume the scaling relations defined in Sijacki et al. (2007), with normalisation constants $D_{\rm bub,0} = 30\ h^{-1}$ kpc, $R_{\rm bub,0} = 50\ h^{-1}$ kpc, $E_{\rm bub,0} = 10^{60}$ erg and $\rho_{\rm ICM,0} = 10^4 h^2 M_\odot$ kpc^{-3} for the bubble distance, radius, energy content and ambient density, respectively.

The remainder of the feedback energy that is not coupled to the surrounding gas by the quasar- or radio-mode goes into radiative electromagnetic feedback, which is approximated as an additional radiation field around the BH superposed with the redshift-dependent ultraviolet background. The full details of this model are given in Vogelsberger et al. (2013).

In Appendix A we present several parametrizations of AGN feedback that were considered during the calibration process to demonstrate how changes to the duty cycle and energetics of the feedback modes impact the $z = 0$ galaxy stellar mass function and the stellar and gas mass fractions of massive haloes. In

particular, we show that the introduction of a quasar-mode duty cycle can greatly suppress the growth of stellar mass in massive galaxies compared to continuous quasar-mode feedback.

2.3 COMPARING TO OBSERVATIONS

The physical properties of a simulation often do not correspond directly to those derived from real observations. This can result from a number of factors, including selection biases, projection effects, instrument systematics and methodology. In this section we will discuss how we can mitigate some of these effects in our comparisons with observations. We also discuss the derivation of quantities which are not intrinsic to the simulation. For example, X-ray luminosity is a complicated combination of continuum and line emission that is not followed in the simulations. We therefore follow a procedure described in Section 2.3.2 in which we create synthetic X-ray spectra in post-processing and analyse them similarly to observations.

2.3.1 *Galaxy stellar mass functions*

In the field of numerical galaxy formation, the galaxy stellar mass function (GSMF) is one of the most important observationally constrained properties that simulations should match. In constructing our GSMF we define galaxies as the self-bound subhaloes identified by SUBFIND and calculate their properties based on the corresponding subhalo catalogue. With this definition, the total stellar mass of a galaxy is the total mass of star particles bound to the subhalo. However, this does not necessarily correspond to the stellar mass of the galaxy as would be measured by an observer. This is because a significant fraction of the stellar mass in massive systems can exist in the form of diffuse intracluster light (ICL), which is difficult to quantify in stellar mass measurements due to its low surface brightness (e.g. Zibetti et al. 2005; Gonzalez et al. 2007; Morishita et al. 2017). Some studies choose to integrate a galaxy's light within a 2-D aperture of a given size (e.g. Li & White 2009) while others integrate Sersic or other fits to the light profiles (e.g. Baldry et al. 2012; Bernardi et al. 2013). This choice can have a significant impact on the derived stellar mass function (see e.g. Bernardi et al. 2013). We follow previous simulation studies in the literature and present our GSMF using multiple definitions of a galaxy's stellar mass: the total stellar mass bound to the corresponding subhalo and two aperture masses. In one case we follow Genel et al. (2014), who consider the stellar mass within a spherical aperture with radius

equal to twice the stellar half-mass radius ($2r_{\star,1/2}$). In the other, we follow Schaye et al. (2015) who define the galaxy stellar mass as the mass within a fixed spherical aperture of radius 30 physical kiloparsec (pkpc). In each case only stellar mass bound to the subhalo is considered.

2.3.2 *X-ray properties*

Galaxy clusters are permeated with diffuse gas heated to temperatures on the order of 10^{7-8}K. This intracluster medium (ICM) emits strongly in X-rays, providing an invaluable means of detecting and studying the properties of clusters. Previous studies have shown that the properties of hot gas derived from X-ray observations can be biased, for example due to complex thermal structure or clumping of the gas (e.g. Mazzotta et al. 2004; Rasia et al. 2005; Nagai et al. 2007a; Khedekar et al. 2013). A reliable comparison between simulations and observations therefore requires actual simulation of the spectral properties of the X-ray emission. We derive X-ray luminosities and spectroscopic temperatures in the following manner.

For a given halo we consider gas within a projected aperture of radius r_{500} centred on the minimum of the gravitational potential. This includes all gas along the line of sight (although for the zoom-in simulations we exclude gas outside the high-resolution region). This is intended to mimic X-ray observations, which measure projected luminosities and temperatures. We exclude cold gas with a temperature less than 3×10^4 K as the lack of molecular cooling in our simulations implies that the temperature of such gas can be significantly overestimated and it should contribute negligibly to the X-ray emission. We also exclude gas above the density threshold required for star formation, as the sub-grid multiphase model for star-forming gas does not reliably predict the thermal properties of the gas.

We project the emission measure of the gas onto temperature bins of width 0.02 keV between 0.01 keV and 24.0 keV, summing up the emission measure in each bin. Using the XSPEC package (Arnaud 1996, version 12.8.0), we then produce a mock X-ray spectrum by summing APEC emission models (Smith et al., 2001) generated for each temperature bin. For simplicity we assume a constant metallicity of 0.3 times the solar value. This ensures that our X-ray analysis remains independent of the metal enrichment in the simulations, which is highly sensitive to the details of the feedback model. Even so, we find no significant difference in the derived X-ray luminosities and temperatures when using the metallicity of the gas in the simulations. The mock spectrum is convolved with the response function of *Chandra* and we adopt a large exposure time of 10^6 seconds so as not to be

limited by photon noise. We then fit the spectrum with a single-temperature APEC model in the 0.5–10 keV energy range with the temperature, metallicity and normalisation of the spectrum left as free parameters. From the best-fitting spectrum we obtain the X-ray luminosity and spectroscopic temperature.

2.3.3 *Halo masses*

As discussed in Section 2.1 and 1.6, the most common method to obtain halo masses for calibrating cluster scaling relations is the X-ray hydrostatic method, which some studies have shown can underestimate the true halo mass. As there is no consensus as to the magnitude of the bias, any comparison between X-ray hydrostatic masses and true halo masses from the simulations must be carefully considered. We make use of the rising number of weak gravitational lensing studies to compare true masses from the simulations with both X-ray hydrostatic masses and halo masses measured with weak lensing. The latter are generally considered to be less biased due to the insensitivity of gravitational lensing to the equilibrium state of the gas or dark matter (for a review, see e.g. Hoekstra et al. 2013; Mandelbaum 2018). On the other hand, lensing mass measurements can possess significantly more scatter due to the effects of cluster substructure, cluster triaxiality and projection effects (e.g. Corless & King 2007; Marian et al. 2010; Meneghetti et al. 2010; Becker & Kravtsov 2011). Since there are also many more observed systems with available X-ray hydrostatic masses than weak lensing masses, especially toward lower halo masses, we choose to compare spherical overdensity masses measured from the simulations to both measures where possible.

2.3.4 *SZ properties*

The thermal Sunyaev–Zel'dovich (SZ) effect has long been recognised as a powerful tool for studying the physics of the ICM and the formation of large-scale structure (Birkinshaw, 1999; Carlstrom et al., 2002). The SZ effect appears as a distortion in the cosmic microwave background (CMB) spectrum, arising from the inverse Compton scattering of CMB photons on energetic electrons in the ICM.

The Compton y parameter (equation 1.12) integrated over solid angles yields an SZ "flux", $Y_{\rm SZ}$, that is proportional to the total thermal energy of the hot gas. Specifically,

$$D_A^2(z)\, Y_{\rm SZ} \equiv \frac{\sigma_T}{m_e c^2} \int P\, dV\,, \tag{2.1}$$

where $D_A(z)$ is the angular diameter distance, σ_T the Thomson cross-section, m_e

the electron rest mass, c the speed of light, and $P = n_e k_B T_e$ the electron pressure, equal to the product of the electron number density, the Boltzmann constant and the electron temperature. In the following we define Y_{500} as the SZ signal integrated within a sphere of radius r_{500} according to equation 2.1.

We compare the SZ signal of our simulated systems with Planck Collaboration XI (2013), who perform a stacking analysis on *Planck* multi-frequency observations of a large sample of Locally Brightest Galaxies (LBGs) selected from the Sloan Digital Sky Survey (SDSS). The selection criteria were designed to maximise the fraction of objects that are the central galaxies of their dark matter haloes. Correspondingly we measure the SZ signal only for central galaxies. Planck Collaboration XI (2013) estimate the mean Y_{500} in a series of stellar mass bins and convert this to a SZ signal-halo mass relation by estimating an "effective" halo mass for each stellar mass bin using a mock sample of LBGs from the semi-analytic galaxy formation simulation of Guo et al. (2011). We also compare to the weak lensing recalibration of the SZ signal-halo mass relation given in Wang et al. (2016), which reduces the dependency of the halo mass estimates on the galaxy formation model and explicitly accounts for uncertainties in both the modelling and the lensing measurements.

Due to the limited angular resolution of *Planck*, the SZ flux is actually measured within a (projected) aperture of radius $5\,r_{500}$. Planck Collaboration XI (2013) convert this measured flux, $Y_{5r_{500}}$, into the flux within a spherical aperture of radius r_{500} by a conversion factor $Y_{500} = Y_{5r_{500}}/1.796$. This factor assumes the spatial template used in their matched filter, the universal pressure profile (Arnaud et al., 2010), and assumes no SZ flux originates from beyond $5\,r_{500}$. In our comparisons we remove the dependency on the assumed pressure distribution by converting the observationally inferred flux reported in Planck Collaboration XI (2013), Y_{500}, back into the actual measured flux, $Y_{5r_{500}}$. This is motivated by the work of Le Brun et al. (2015) who show that Y_{500} is highly sensitive to the assumed spatial template. Le Brun et al. (2015) generate synthetic SZ maps from a cosmological hydrodynamical simulation from the cosmo-OWLS project that reproduces a range of global SZ, X-ray and optical properties of local groups and clusters (Le Brun et al., 2014). Applying the same tools and assumptions as *Planck*, Le Brun et al. (2015) show that the inferred flux, Y_{500}, is biased high by a factor ~ 2 at $M_{500} = 2.6 \times 10^{13} M_{\odot}$. This bias increases with decreasing halo mass, reaching nearly an order of magnitude overestimate below $\sim 10^{13} M_{\odot}$. The vast majority of the bias is due to the assumption of a fixed spatial template, which becomes an increasingly worse description of the hot gas in low mass haloes. An analysis of the ratio $Y_{5r_{500}}/Y_{500}$ for FABLE systems seems to support this result. The median

$Y_{5r_{500}}/Y_{500}$ for the seven FABLE clusters with $M_{500} > 5 \times 10^{14} M_\odot$ at $z = 0$ is 1.73, consistent with the constant value of 1.796 used in the *Planck* analysis. However, the median ratio increases towards lower mass haloes, reaching $Y_{5r_{500}}/Y_{500} = 3.77$ at $10^{13} M_\odot \leq M_{500} \leq 5 \times 10^{13} M_\odot$. This is a factor of two larger than the value assumed by Planck Collaboration XI (2013) and implies that their inferred Y_{500} is biased high on galaxy group scales.

The SZ flux as measured by *Planck* for individual, low-mass systems may be biased due to source confusion (i.e. hot gas along the line-of-sight). Indeed, our tests have shown that the flux of individual low-mass systems can be significantly boosted by hot gas which overlaps with them in projection. However, Planck Collaboration XI (2013) estimate the mean SZ flux from stacking a large number of systems in mass bins, which can remain unbiased by source confusion. Indeed, Le Brun et al. (2015) demonstrate that the mean SZ flux is not significantly biased by uncorrelated confusion across the whole mass range of the Planck Collaboration XI (2013) stacking analysis. For this reason we integrate the SZ flux of a galaxy within a spherical aperture of radius $5\,r_{500}$ and note that this may underestimate the observed flux if correlated or uncorrelated source confusion leads to a significant bias in the observations.

2.3.5 *ICM profiles*

Radial profiles of the intracluster medium characterize its distribution and thermodynamic history. The effects of non-gravitational processes such as AGN feedback cause ICM profiles to deviate from the self-similar relations predicted in the absence of such processes (e.g. Voit et al. 2005). A comparison of simulated profiles to observed ones is therefore a useful test of non-gravitational physics.

In X-ray observations, gas density profiles are derived from background-subtracted surface brightness profiles and gas temperature profiles are obtained from extracted spectra. Since these observations are sensitive only to hot X-ray emitting gas, in calculating the radial profiles of our simulated ICM we apply the same temperature–density cuts as used in our X-ray analysis described in Section 2.3.2. The gas is then divided into concentric spherical shells with logarithmically-spaced radii centred on the minimum of the gravitational potential of the halo. For each radial bin we calculate the volume-weighted electron number density, n_e, and the mass-weighted temperature, T. The former is defined as the total electron number divided by the volume of the bin. Since we may exclude some gas cells from the bin, we correct the bin volume by the ratio of the total volume of cells that are not excluded to the total volume of all cells in the bin.

2.4 THE GALAXY POPULATION

2.4.1 *Galaxy stellar mass function at $z = 0$*

Figure 2.2 shows the $z = 0$ galaxy stellar mass function (GSMF) for all galaxies in the $(40\ h^{-1}\ \mathrm{Mpc})^3$ simulation volume. We plot the GSMF for three different definitions of a galaxy's stellar mass, as discussed in Section 2.3.1: these are the total bound stellar mass, the mass within twice the stellar half-mass radius ($2\,r_{\star,1/2}$) and the mass within a radius of 30 pkpc. We compare with the Illustris and EAGLE simulations (grey lines) and observational estimates of the GSMF from Li & White (2009), Baldry et al. (2012), Bernardi et al. (2013) and D'Souza et al. (2015) (symbols with error bars). We plot the $z = 0$ GSMF from Illustris for both the total bound mass and the mass within $2\,r_{\star,1/2}$ (Genel et al., 2014) and the $z = 0.1$ GSMF of EAGLE, which uses the mass within a 30 pkpc aperture (Schaye et al., 2015).

At the low mass end of the GSMF, FABLE is in excellent agreement with the observational data. This contrasts with Illustris, which significantly overestimates the abundance of galaxies below the knee ($M_\star \lesssim 10^{10} M_\odot$). While some of this improvement is owed to changes to the stellar feedback model, we have determined that the dominant cause is a difference in the gravitational softening compared to Illustris. We have chosen somewhat larger softening lengths compared to Illustris and find that this results in fewer low-mass galaxies. Somewhat surprisingly, the difference is not only present near the resolution limit but extends almost two decades in stellar mass. We have tested this explicitly by performing a simulation identical to the FABLE periodic box described in Section 2.2.1 but with a gravitational softening length that is approximately 2.5 times smaller. Quantitatively, the effect is a systematic increase of ~ 0.2 dex in the mass function for stellar masses ranging from near the resolution limit ($\sim 10^8 M_\odot$) to just below the knee ($\sim 10^{10} M_\odot$). This offset is already present at $z = 8$ and persists until $z = 0$. As there is no general consensus regarding the choice of gravitational softening lengths, our improved agreement at the low mass end of the GSMF is largely serendipitous. We also note that, like Illustris, we are unlikely to be fully converged with respect to resolution.

The knee of the GSMF is in agreement with the data, although it is slightly lower than in Illustris. The difference is largely the result of changes to the stellar feedback model. As discussed in Section 2.2.3, we have modified the stellar feedback model of Illustris by assigning one-third of the galactic wind energy as thermal. We find that this reduces the abundance of galaxies mostly around the knee of the $z = 0$ GSMF, leaving lower masses ($M_\star \lesssim 10^{10} M_\odot$) relatively unaf-

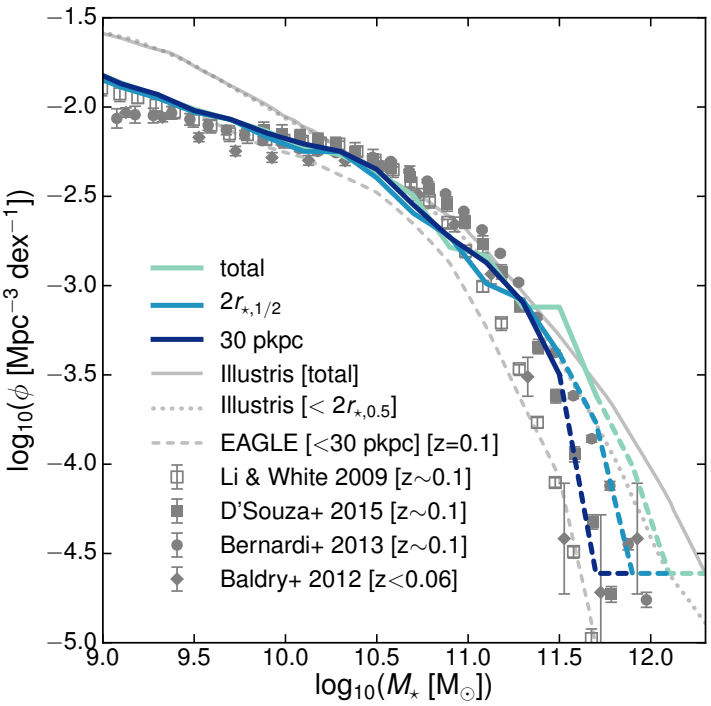


Figure 2.2: The FABLE galaxy stellar mass function at $z = 0$ for different definitions of galaxy stellar mass (lines) in comparison to observations (symbols with error bars). We consider three definitions of a galaxy's stellar mass: the total mass of all star particles bound to the subhalo, those within twice the stellar half-mass radius and those within 30 pkpc. Lines become dashed at the high-mass end when there are fewer than 10 objects per 0.2 dex stellar mass bin. Data points show measurements with 1-sigma error bars from Li & White (2009), D'Souza et al. (2015), Bernardi et al. (2013) and Baldry et al. (2012). The grey solid and dotted lines are the $z = 0$ GSMFs of Illustris (Genel et al., 2014) using the total bound stellar mass of each galaxy and the mass within twice the stellar half-mass radius, respectively. The grey dashed line is the $z = 0.1$ GSMF of EAGLE (Schaye et al., 2015) using the stellar mass within a spherical 30 pkpc aperture. All mass functions assume a Chabrier (2003) IMF.

fected. Physically, star formation is suppressed because warmer winds increase the buoyancy of the gas outflow thereby reducing the rate at which gas is recycled back onto the galaxy.

The high mass end of the total and $2\,r_{\star,1/2}$ GSMFs are very similar to those of Illustris, with a slight reduction in the abundance of the most massive galaxies. This is encouraging given that we have greatly reduced the burstiness of the radio-mode of AGN feedback compared to Illustris and therefore reduced its ability to suppress star formation in massive haloes. Instead we have managed to efficiently suppress star formation in massive galaxies through the use of a duty cycle for the quasar-mode of feedback. In Appendix A we show how this affects the high mass end of the $z = 0$ GSMF and the gas fractions of massive haloes compared

with continuous quasar-mode feedback.

The GSMF of EAGLE was calibrated to reproduce the mass function derived by Li & White (2009) for a complete spectroscopic sample from the Sloan Digital Sky Survey (SDSS) (open symbols in Fig. 2.2). Li & White (2009) measure galaxy flux within a projected Petrosian aperture, which Schaye et al. (2015) show yields a GSMF similar to a spherical aperture of radius 30 pkpc for galaxies in EAGLE. At low masses, $M_\star \lesssim 10^{10} M_\odot$, our 30 pkpc aperture mass function is almost identical to that of EAGLE. Around the knee of the GSMF ($M_\star \lesssim 10^{11} M_\odot$) we are actually in better agreement with Li & White (2009) compared with EAGLE, which somewhat underestimates the GSMF there. At larger masses we overestimate Li & White (2009) by ~ 0.3 dex, although the difference may be exaggerated as Bernardi et al. (2017) argue that Li & White (2009) use mass-to-light ratios that are biased low for massive galaxies.

At the high mass end of the GSMF there is significant variation between different observational studies. Some of the dominant causes for this variation are discussed in Bernardi et al. (2017). In particular, the observed GSMF depends on how the total light associated with a galaxy is determined. The potential impact of this effect is apparent from the simulated total and aperture mass functions shown in Fig. 2.2, which differ by ~ 0.3 dex on the vertical axis at the high mass end. In addition, there is some freedom in the stellar population modelling used to estimate the mass associated with the total stellar light (i.e. the mass-to-light ratio).

The GSMF derived from the total bound stellar mass of galaxies slightly exceeds observational constraints at the high mass end, suggesting that the FABLE simulations produce slightly too many massive galaxies. The degree to which this is true depends on what fraction of the total mass in massive galaxies is accounted for in observations. In particular, a significant fraction (~ 30 per cent) of the total stellar mass in our massive galaxies is contained in the ICL, the diffuse nature of which makes it difficult to quantify from observations. Studies which measure the galaxy flux within a particular aperture will typically exclude the majority of the ICL associated with massive galaxies. For example, the Petrosian aperture used by Li & White (2009) is known to significantly underestimate the flux of galaxies with extended surface brightness profiles (Blanton et al., 2001; Graham et al., 2005; Bernardi et al., 2010, 2013). More recent studies such as Baldry et al. (2012), Bernardi et al. (2013) and D'Souza et al. (2015) attempt to measure a better estimate of the total flux of galaxies by integrating models fit to their surface brightness distributions. Baldry et al. (2012) fit Sersic profiles to $z < 0.06$ galaxies from the Galaxy And Mass Assembly (GAMA) survey while Bernardi et al. (2013)

fit Sersic-exponential models to a magnitude-limited sample of SDSS galaxies. D'Souza et al. (2015) use the same sample as Li & White (2009) but integrate the galaxy flux from exponential or de Vaucouleurs profile fits and derive flux corrections from stacked SDSS images, which provide a more accurate measurement of the total amount of light owing to the increased signal-to-noise ratio. The extra light returned by this method compared to Li & White (2009) results in a larger abundance of massive galaxies, as evident in Fig. 2.2. These model profiles can only be fit to the central, high signal-to-noise regions of galaxies and must make assumptions about the outer regions of the galaxy profile. Which profile is the most appropriate at the high mass end is still debated and can lead to a significant bias in the total estimated flux and resultant stellar mass estimate (see e.g. discussion in Bernardi et al. 2013).

Using an aperture of radius $2\,r_{\star,1/2}$, the FABLE GSMF is in very good agreement with Bernardi et al. (2013) at the high mass end but is overestimated compared to D'Souza et al. (2015) and Baldry et al. (2012). Bernardi et al. (2017) show that the Sersic-exponential fits used by Bernardi et al. (2013) return galaxy fluxes similar to the corrected fluxes of D'Souza et al. (2015). Similarly, Bernardi et al. (2013) show that their luminosity function is in good agreement with that used in Baldry et al. (2012). This implies that the difference between these three studies at the massive end of the GSMF is due to differences in their assumed mass-to-light ratios rather than their methods for estimating galaxy flux. Bernardi et al. (2017) state that the Bernardi et al. (2013) model is oversimplified and may overestimate the mass-to-light ratio. On the other hand, the stellar population modelling used by D'Souza et al. (2015) results in a mass function that is ~ 0.3 dex lower on the vertical axes above $10^{11.5} M_{\odot}$ compared to more recent estimates of the mass-to-light ratio based on the same IMF (Bernardi et al., 2017). Given that there is no consensus as to the best approach to stellar population modelling, in addition to the uncertainty in how to fit the light profiles of massive galaxies, there is arguably little point in tuning the simulated galaxy stellar mass function to a specific dataset. Overall, we are satisifed that the difference between the simulated and observed GSMFs is similar to the variation between different observational studies.

2.4.2 *Galaxy stellar mass function at $z \leq 3$*

In Fig. 2.3 we plot the GSMF at $z \leq 3$ in comparison to observational data from Muzzin et al. (2013) and Ilbert et al. (2013), two independent estimates both based on UltraVISTA DR1, and Tomczak et al. (2014), for the FourStar Galaxy Evolution Survey (ZFOURGE).

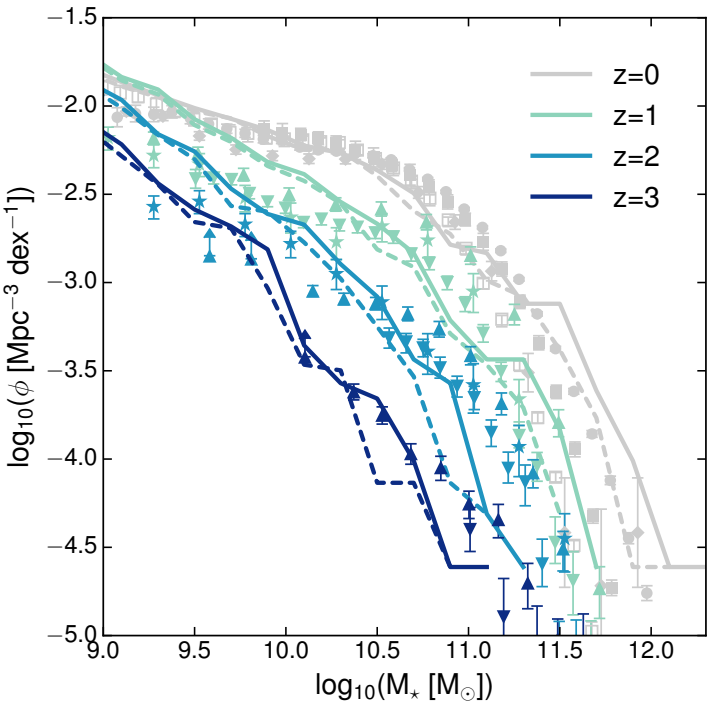


Figure 2.3: The galaxy stellar mass function at redshifts $0 \leq z \leq 3$ (lines) compared to observations (symbols with error bars). Two definitions are used for a simulated galaxy's stellar mass: all stellar mass bound to the subhalo (solid lines) and bound stellar mass within twice the stellar half-mass radius (dashed lines). The $z = 0$ data are the same as in Fig. 2.2. At $z = 1$ we compare to observed GSMFs in the redshift ranges $1.0 \leq z < 1.5$, $0.8 < z < 1.1$ and $1.0 < z < 1.25$ for Muzzin et al. (2013) (downward triangles), Ilbert et al. (2013) (upward triangles) and Tomczak et al. (2014) (stars), respectively. At $z = 2$ and $z = 3$ we compare with the GSMFs for redshift ranges $2.0 < z < 2.5$ and $3.0 < z < 3.5$, respectively. Only stellar mass bins above the mass completeness limit are plotted in each case. Stellar masses have been converted to a Chabrier (2003) IMF where necessary by subtracting 0.25 dex or 0.05 dex for a Salpeter (1955) or Kroupa (2001) IMF, respectively.

We continue to have good agreement with the data beyond $z = 0$. Although the FABLE model for AGN feedback has been calibrated to match the $z = 0$ GSMF, the agreement is not guaranteed at higher redshifts. The high mass end of the GSMF is in good agreement with the data at each of the redshifts shown, except for a slight underestimate at $z = 2$. This may be due to small number statistics imposed by our finite box size, as the two highest occupied mass bins at $z = 2$ contain only one or two galaxies.

At $z \geq 1$ the low mass end of the GSMF is somewhat overestimated, although this is not entirely unexpected given that this was also the case in Illustris. We have slightly altered the stellar feedback model of Illustris by implementing thermal winds, however this largely affects galaxies around the knee of the GSMF rather than low mass galaxies and any significant changes are limited to $z < 2$.

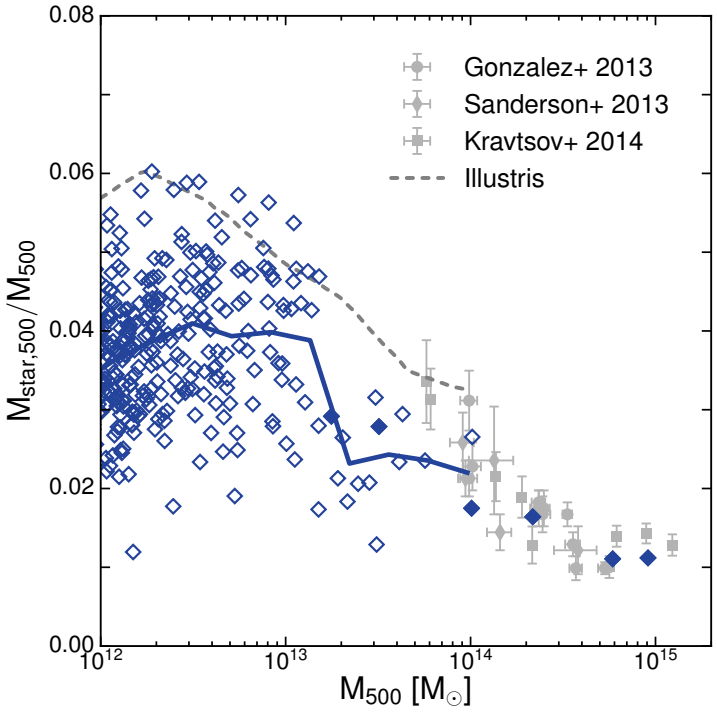


Figure 2.4: Stellar mass fraction within r_{500} as a function of halo mass at $z = 0$. Open diamonds show haloes from the $40\,h^{-1}$ Mpc periodic box while filled diamonds show the main halo of each zoom-in simulation. The solid line shows the mean relation in bins of halo mass for comparison with the mean relation from Illustris at $z = 0$ (grey dashed line). Symbols with error bars show $z \simeq 0$ observations from Gonzalez et al. (2013), Sanderson et al. (2013) and Kravtsov et al. (2018). Following Chiu et al. (2016b), the stellar mass measurements of Gonzalez et al. (2013) and Sanderson et al. (2013) have been reduced by 24 per cent to ensure all stellar masses are appropriate for a Chabrier IMF.

2.5 GLOBAL GROUP AND CLUSTER PROPERTIES

2.5.1 *Stellar mass fractions*

In Fig. 2.4 we plot the stellar mass fraction within r_{500} as a function of halo mass at $z = 0$. We consider all FoF haloes in the periodic volume (open diamonds), plus the main FoF halo in each of the zoom-in simulations (filled diamonds). We do not discriminate between stars in satellite galaxies, the brightest central galaxy (BCG) or the ICL because the distinction between BCG and ICL is not well defined. We therefore compare to studies which take into account contributions from all three components (Gonzalez et al., 2013; Sanderson et al., 2013; Kravtsov et al., 2018). The grey dashed line is the mean relation from the Illustris simulation (Genel et al., 2014).

The FABLE simulations are a very good match to observed stellar mass fractions across a wide range of halo masses from galaxy groups (a few times $10^{13} M_\odot$)

to high mass clusters ($\sim 10^{15} M_\odot$). Although our AGN feedback model was tuned to reproduce the $z = 0$ GSMF, the cosmological volume used in the tuning process contains only one halo with $M_{500} \sim 10^{14} M_\odot$. It is therefore reassuring that our model yields a realistic buildup of stellar mass even in dense cluster environments. We reproduce the observed trend with halo mass, with an increase in the stellar fraction toward lower mass systems. The relationship between stellar fraction and halo mass is similar between FABLE and Illustris, with an offset in the normalisation. The difference in normalisation at $M_{500} \gtrsim 10^{13} M_\odot$ is due to more efficient suppression of star formation by AGN feedback in FABLE, consistent with the offset in the GSMFs. The offset at $M_{500} \lesssim 10^{13} M_\odot$ is partly due to our choice of larger gravitational softening lengths compared with Illustris and partly our change to the stellar feedback model.

We caution that the observations plotted in Fig. 2.4 derive halo masses from X-ray data, which could potentially underestimate the true mass (see Section 2.3.3 and 1.6). Correcting for such a bias (if present) would shift the observational data to lower stellar fractions, away from our relation. On the other hand, observations may underestimate the stellar mass of the BCG and associated ICL, which can contribute $\gtrsim 50$ per cent of the total stellar mass both observationally (e.g. Gonzalez et al. 2013) and in our simulations. The characteristically diffuse emission of the ICL makes it particularly difficult to quantify and thus a non-negligible fraction of a cluster's total stellar content may be missed.

2.5.2 *Gas mass fractions*

In Fig. 2.5 we plot the total gas mass fraction within r_{500} as a function of halo mass at $z = 0$. The solid line shows the mean relation in bins of halo mass. We compare to a range of observational data (symbols with error bars) as well as the mean relation in Illustris (grey dashed line). Since observed gas fractions are obtained from X-ray observations, we exclude cold gas and star-forming gas that is followed only with a simplified multiphase model, in accordance with our simulated X-ray spectra (see Section 2.3.2). For completeness, we also show the mean gas fraction–halo mass relation when all gas is included (dashed line). The difference becomes significant below $\sim 10^{13} M_\odot$, however there is very little available X-ray data at these masses.

We have calibrated the FABLE AGN feedback model to observed gas fractions only for haloes in the $40\,h^{-1}$ Mpc periodic box, which are represented by the open diamonds in Fig. 2.5. Although the model was not calibrated to cluster scales ($\gtrsim 10^{14} M_\odot$), the zoom-in simulations (filled diamonds) are in good agreement

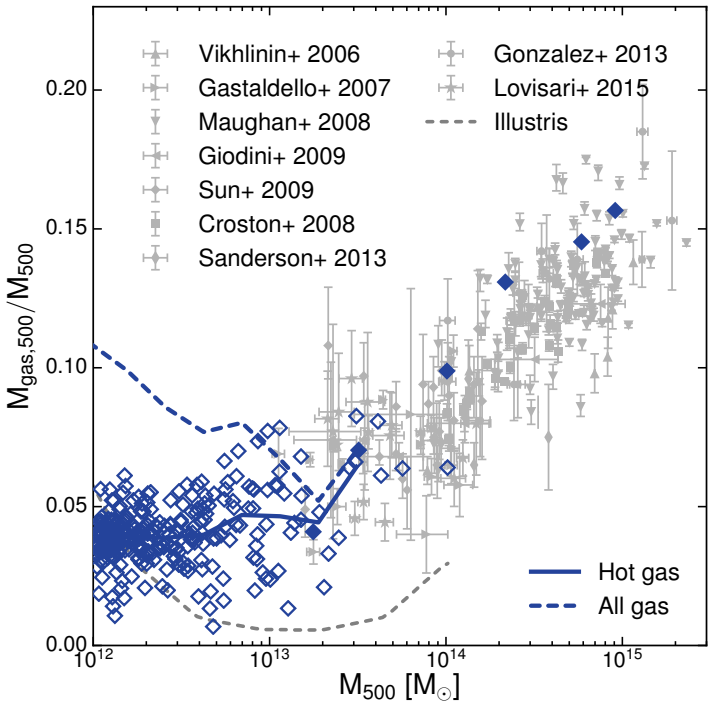


Figure 2.5: Gas mass fraction within r_{500} as a function of halo mass at $z = 0$. Marker styles are the same as in Fig. 2.4. The solid line shows the mean relation of these points in halo mass bins. Cold and multiphase gas has been excluded as described in Section 2.3.2. The dashed line shows the mean relation when all gas is included. The grey dashed line is the mean relation from Illustris at $z = 0$ (Genel et al., 2014). Symbols with error bars show $z \simeq 0$ observations from Vikhlinin et al. (2006), Gastaldello et al. (2007), Maughan et al. (2008), Giodini et al. (2009), Sun et al. (2009), Croston et al. (2008), Sanderson et al. (2013), Gonzalez et al. (2013) and Lovisari et al. (2015).

with the observations even in massive clusters ($\approx 10^{15} M_\odot$). This was certainly not guaranteed, as the much deeper potentials of clusters make it more difficult for AGN feedback to eject gas beyond r_{500}. Indeed, the mean gas fractions of our simulated clusters lie slightly above the mean observed relation, which suggests that the AGN feedback may need to be slightly more efficient at these high masses. We note that this offset would be exacerbated in the case of a significant X-ray mass bias, which would shift the observed data away from our relation.

At $M_{500} \approx 10^{13} - 10^{14} M_\odot$ the mean gas mass fraction in FABLE is significantly higher than in Illustris. In the Illustris model, radio-mode AGN feedback ejected large gas masses out of massive haloes, resulting in significantly underestimated gas fractions. In our updated model, radio-mode feedback events are more frequent but less energetic than in Illustris, and are therefore less able to eject gas from massive haloes. In Appendix A we show that radio-mode feedback in FABLE is able to lower gas fractions in massive haloes but that it is also assisted by

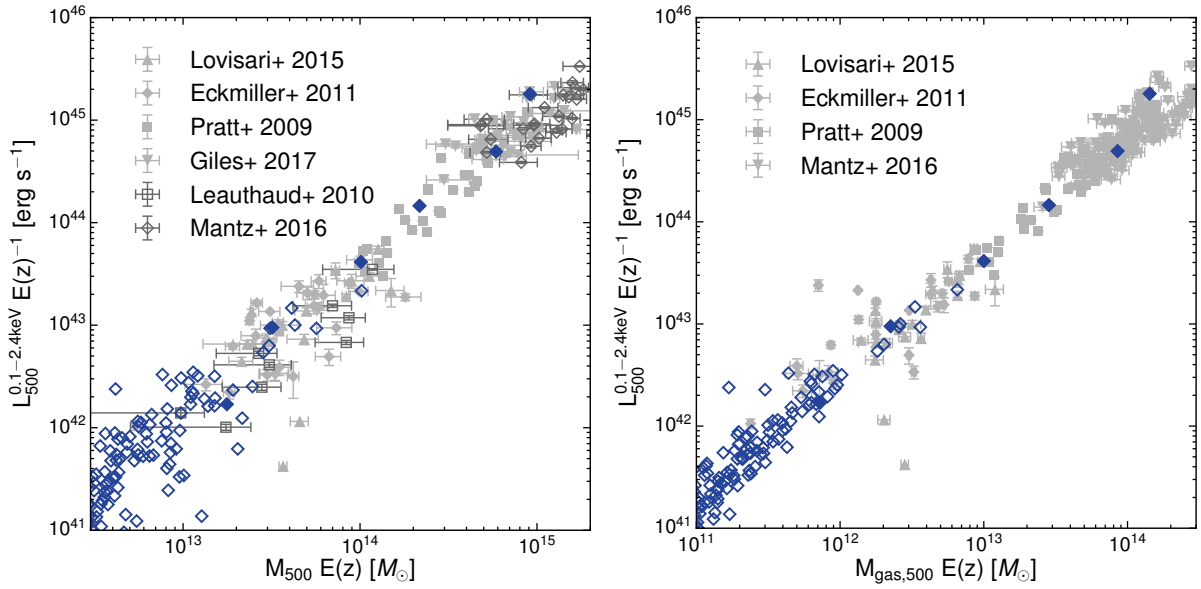


Figure 2.6: X-ray Luminosity in the 0.1–2.4 keV band as a function of total mass (left) and gas mass (right) at $z = 0$. Marker styles are the same as in Fig. 2.4. In the left hand panel, light grey symbols with error bars represent observational data for which total masses were estimated from X-ray observations assuming hydrostatic equilibrium (Lovisari et al., 2015; Eckmiller et al., 2011; Pratt et al., 2009; Giles et al., 2017). Dark grey symbols with error bars show observational data for which total masses were estimated from weak gravitational lensing (Leauthaud et al., 2010; Mantz et al., 2016b). In the right hand panel we compare to X-ray luminosities and gas masses from Lovisari et al. (2015), Eckmiller et al. (2011), Pratt et al. (2009) and Mantz et al. (2016b).

the quasar-mode, whose periodic rather than continuous feedback reduces gas fractions by suppressing the accumulation of gas at early times.

2.5.3 *X-ray Luminosity-Mass Relation*

Here we present X-ray luminosity as a function of halo mass and gas mass in the FABLE simulations in comparison with observations. X-ray luminosities are calculated as described in Section 2.3.2 in one of two bands: a soft X-ray band in the range 0.1–2.4 keV and a bolometric band in the range 0.01–100 keV. We scale luminosity by $E(z)^{-1}$ and mass by $E(z)$, where $E(z) = \sqrt{\Omega_{\rm M}(1+z)^3 + \Omega_\Lambda}$ describes the redshift evolution of the Hubble parameter (see derivation in Section 1.2.1.2). Halo masses are those measured directly from the simulation. The gas mass is the total mass of gas included in the creation of the synthetic X-ray spectrum, i.e., after excluding cold and multiphase gas (see Section 2.3.2).

In Fig. 2.6 we plot the $L_{500} - M_{500}$ and $L_{500} - M_{\rm gas}$ relations in the soft X-ray band. At group-scales we compare to Lovisari et al. (2015), a complete X-ray

selected sample of 20 galaxy groups observed with *XMM-Newton*, and Eckmiller et al. (2011), a sample of 26 groups observed with *Chandra*. At the high mass end we compare to Pratt et al. (2009) who investigate the luminosity scaling relations of 31 local ($z < 0.2$) clusters from the Representative XMM-Newton Cluster Structure Survey (REXCESS; Böhringer et al. 2007). REXCESS halo masses were estimated iteratively from the $M_{500} - Y_{\rm X}$ relation of Arnaud et al. (2007), where $Y_{\rm X}$ is the product of the gas mass and core-excised temperature, and gas masses are taken from Croston et al. (2008). For the $L_{500} - M_{500}$ relation we supplement the high mass end with data from Giles et al. (2017), a complete sample of 34 galaxy clusters at $0.15 \leq z \leq 0.3$ observed with Chandra. Each of these studies uses X-ray hydrostatic mass estimates.

We also compare to the weak lensing calibrated $L_{500} - M_{500}$ relations of Leauthaud et al. (2010) and Mantz et al. (2016b). Leauthaud et al. (2010) measure mean halo masses via stacked weak gravitational lensing for a sample of 206 X-ray selected groups in the COSMOS field. We have rescaled their mean halo mass, $\langle M_{200} \rangle$, to the mass within r_{500} using the best-fit NFW profile of each mass bin. Mantz et al. (2016b) measure X-ray luminosities from *Chandra* and *ROSAT* data for 27 clusters with weak lensing mass estimates as part of the *Weighing The Giants* project (Applegate et al., 2014; Kelly et al., 2014; von der Linden et al., 2014a). We also plot the $L_{500} - M_{\rm gas}$ measurements of Mantz et al. (2016b) for their full sample of 139 clusters.

Compared with observed $L_{500} - M_{500}$ relations based on X-ray hydrostatic masses, the FABLE simulations are in excellent agreement over the full mass range from groups to massive clusters. However, there is tentative evidence that the predicted relation is overestimated compared with observations based on weak lensing masses. This may point to the existence of an X-ray mass bias, which would manifest itself as a systematic difference between the $L_{500} - M_{500}$ relations derived from weak lensing masses and those derived from X-ray hydrostatic masses. If a significant X-ray mass bias is indeed present, the observed halo gas mass fractions shown in Fig. 2.5 would be shifted to lower values, away from our relation. In this case, the FABLE clusters would be too gas-rich and as a result their X-ray luminosities would be too high relative to the weak-lensing calibrated $L_{500} - M_{500}$ (assuming weak lensing masses are less biased). Given that X-ray luminosity is very sensitive to the density distribution of the X-ray emitting gas, the luminosities of the simulated systems could be biased by high-density clumps with high X-ray emission. However, the excellent agreement with the observed $L_{500} - M_{\rm gas}$ relation in the right hand panel of Fig. 2.6 suggests that this is not the case, i.e., that the gas content of our simulated haloes has a realistic clumping

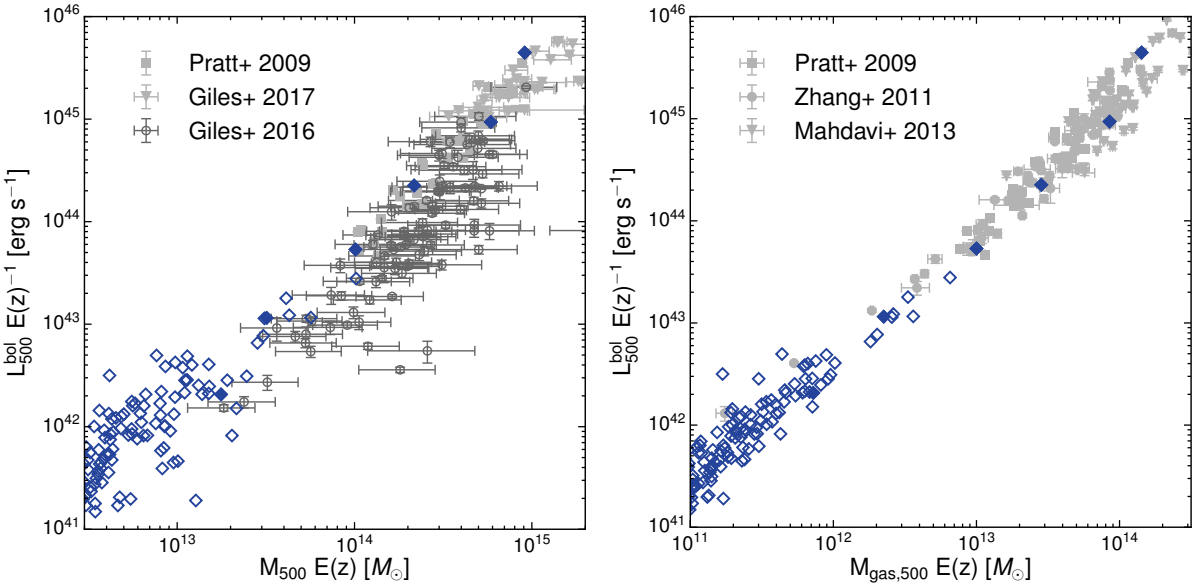


Figure 2.7: Bolometric (0.01–100 keV) X-ray luminosity as a function of total mass (left) and gas mass (right) at $z = 0$. In the left hand panel, light grey symbols with error bars represent observational data based on X-ray hydrostatic masses (Pratt et al., 2009; Giles et al., 2017). Dark grey symbols with error bars are the XXL-100-GC clusters (Giles et al., 2016) for which total masses were estimated from the internally calibrated weak lensing mass–temperature relation presented in Lieu et al. (2016). In the right hand panel we compare with data from Pratt et al. (2009), Zhang et al. (2011) and Mahdavi et al. (2013).

factor.

In Fig. 2.7 we show the $L_{500}-M_{500}$ and $L_{500}-M_{\rm gas}$ relations using the bolometric X-ray luminosity. For the $L_{500} - M_{500}$ relation we compare to REXCESS (Pratt et al., 2009) and Giles et al. (2017), both of whom use X-ray hydrostatic masses. We also compare with weak lensing calibrated data from Giles et al. (2016) for the XXL-100-GC sample, which consists of the 100 brightest clusters in the XXL survey (Pacaud et al., 2016). Halo masses for the XXL-100-GC are estimated from the weak lensing mass–temperature relation presented in Lieu et al. (2016). For $L_{500} - M_{\rm gas}$ we complement the low mass end with data from Zhang et al. (2011), who analyse 62 galaxy clusters in the HIFLUGCS sample with *XMM-Newton* and *ROSAT* data. At the high mass end of $L_{500} - M_{\rm gas}$ we compare with Mahdavi et al. (2013), a sample of 50 clusters with weak lensing mass estimates and X-ray data from *Chandra* or *XMM-Newton*. Note that the luminosity and gas mass for this sample are measured within r_{500} as derived from weak lensing.

Again we achieve excellent agreement with the relation based on X-ray hydrostatic masses. However, as was hinted at in the soft band $L_{500} - M_{500}$ relation, there appears to be a systematic offset between observed bolometric $L_{500} - M_{500}$

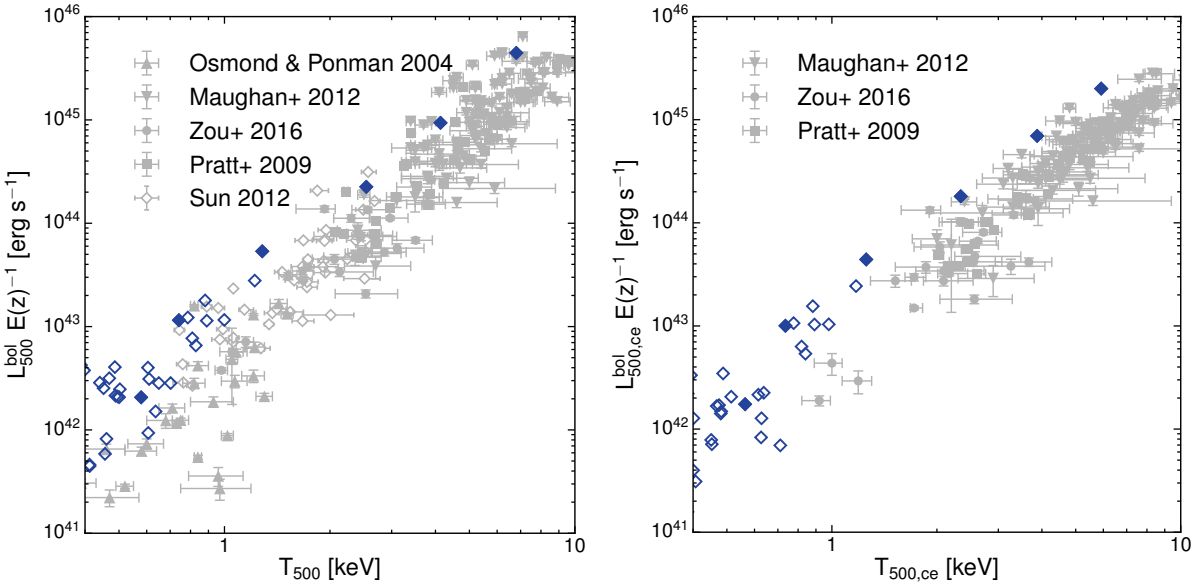


Figure 2.8: Bolometric (0.01–100 keV) X-ray luminosity as a function of spectroscopic temperature measured in the $(0–1)\ r_{500}$ aperture (left) and the $(0.15–1)\ r_{500}$ aperture (right) at $z = 0$. Data from several X-ray studies are shown for comparison (Osmond & Ponman, 2004; Maughan et al., 2012; Zou et al., 2016; Pratt et al., 2009; Sun, 2012).

relations based on weak lensing masses and those based on X-ray hydrostatic masses. Under the assumption that weak lensing masses are less biased, this suggests that the X-ray luminosities of the FABLE groups and clusters may be slightly too high for a given halo mass. This conclusion is, however, complicated by the relatively large measurement uncertainties and overall scatter in weak lensing mass estimates compared to X-ray masses (see e.g. Meneghetti et al. 2010) and the relatively poor sampling of the $L_{500} - M_{500}$ plane by weak lensing data.

2.5.4 *X-ray Luminosity-Temperature relation*

In Fig. 2.8 we plot bolometric X-ray luminosity as a function of X-ray spectroscopic temperature. In the left and right hand panels we consider luminosities and temperatures measured with and without the cluster core, respectively. We define a core radius of $0.15\ r_{500}$ as used in the comparison studies.

We compare against data from Osmond & Ponman (2004), Maughan et al. (2012), Zou et al. (2016), REXCESS (Pratt et al., 2009) and Sun (2012). The Osmond & Ponman (2004) sample contains 35 systems with group-scale X-ray emission observed with *ROSAT*. Although not statistically representative, the sample corresponds to some of the lowest temperature systems that have been observed. The data from Maughan et al. (2012), Zou et al. (2016) and Sun (2012) each consist of

groups and clusters observed with *Chandra*. The Maughan et al. (2012) sample consists of 114 clusters originally described in Maughan et al. (2008), the Zou et al. (2016) sample is statistically complete and contains 23 groups and low-mass clusters from the 400d survey (Burenin et al., 2007), and Sun (2012) presents the X-ray luminosity–temperature relation of the Sun et al. (2009) sample of 43 galaxy groups with which we compare radial profiles of the ICM in Section 2.6.

We find that the FABLE groups and clusters lie on the upper end of the scatter in the observations. This is true whether or not the cluster core is excised, which implies that our X-ray luminosities or temperatures are not biased by, for example, an overabundance of dense cool cores. At $T_{500} \sim 1$ keV we are actually in good agreement with the Sun (2012) data, however, we expect that their sample is biased towards high X-ray luminosities due to their selection criteria, which require that the group emission can be traced to at least r_{2500} ($\approx 0.47\, r_{500}$; Sun et al. 2009). Indeed, their average luminosity at $T_{500} \sim 1$ keV is notably higher than the Osmond & Ponman (2004) and Zou et al. (2016) samples. Similarly, Le Brun et al. (2014) find that the Sun et al. (2009) sample has a significantly higher mean X-ray luminosity for groups with masses $M_{500} \sim 10^{13-13.5} M_{\odot}$ compared with the galaxy group studies of Osmond & Ponman (2004), Rozo et al. (2009) and Leauthaud et al. (2010).

It is not clear whether the discrepancy between the simulation prediction and the observed relations is a result of overestimated total luminosities, underestimated global temperatures, or a combination of the two. The former explanation may be consistent with our comparison to weak lensing studies of the $L_{500} - M_{500}$ relation (Figs. 2.6 and 2.7), for which there is some evidence that our luminosities are too high for a given halo mass. However, the poor sampling and relatively large scatter of the weak lensing measurements means that we cannot rule out the possibility that the gas in our groups and clusters possess too-low average temperatures. In the following section we therefore investigate the role of temperature using the total mass–temperature relation.

2.5.5 *Mass-Temperature relation*

In Fig. 2.9 we show the total mass as a function of spectroscopic temperature in FABLE compared to observational samples that use either X-ray hydrostatic masses (left panel) or weak lensing masses (right panel).

For the X-ray mass comparison we show data from Lovisari et al. (2015), Eckmiller et al. (2011), REXCESS (Pratt et al., 2009) and Mahdavi et al. (2013). We calculate the spectroscopic temperature within r_{500}, appropriate for REXCESS

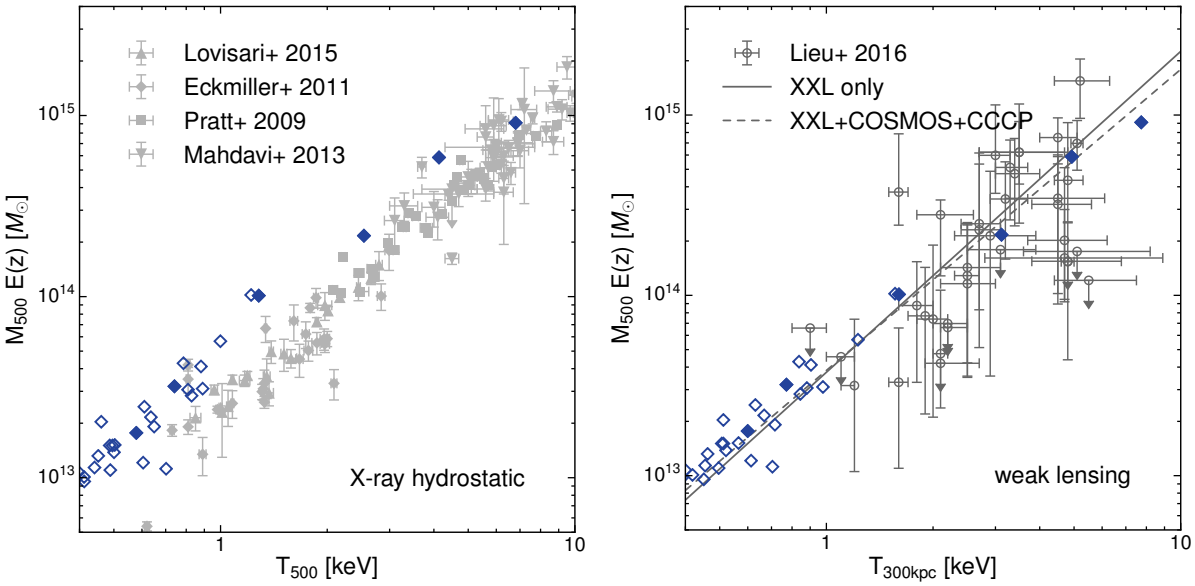


Figure 2.9: Total mass as a function of spectroscopic temperature compared to observations based on X-ray hydrostatic masses (left) and weak lensing masses (right). In the left hand panel temperatures are measured within a projected aperture of radius r_{500}. Symbols with error bars show data from Lovisari et al. (2015), Eckmiller et al. (2011), Pratt et al. (2009) and Mahdavi et al. (2013). In the right hand panel we compare to Lieu et al. (2016) whom measure weak-lensing masses for a sample of the brightest clusters in the XXL survey (symbols with error bars). We mimic Lieu et al. (2016) by measuring temperatures within a projected aperture of radius 300 pkpc. The solid line is the best-fit to the XXL data and the dashed line is the best-fit to an extended sample including additional groups and clusters from COSMOS (Kettula et al., 2013) and CCCP (Hoekstra et al., 2015).

and Mahdavi et al. (2013), but caution that the data of Lovisari et al. (2015) and Eckmiller et al. (2011), which populate group-scale masses, cover only a fraction of r_{500} and this fraction is not consistent between different systems.

For the weak lensing mass comparison we show the mass–temperature relation derived by Lieu et al. (2016) for 38 clusters from the XXL-100-GC sample that lie within the footprint of the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS), using the CFHTLenS shear catalogue for the mass measurements (Heymans et al., 2012; Erben et al., 2013). As in Lieu et al. (2016), we measure spectroscopic temperatures within a projected aperture of radius 300 pkpc. We also show the best-fitting $M_{500} - T_{300\mathrm{kpc}}$ relation from Lieu et al. (2016) for an extended sample including 10 galaxy groups from COSMOS (Kettula et al., 2013) and 48 massive clusters from the Canadian Cluster Comparison Project (CCCP; Hoekstra et al. 2015).

We find that the FABLE groups and clusters lie systematically above the $M_{500} - T_{500}$ relation derived from X-ray masses (left hand panel of Fig. 2.9). As the size

of the offset is similar to the mean offset between the simulated and observed $L_{500} - T_{500}$ relations, both discrepancies could be the result of systematically low average temperatures. On the other hand, the FABLE systems are a good match to the weak lensing calibrated $M_{500} - T_{300\mathrm{kpc}}$ relation from Lieu et al. (2016) (right hand panel of Fig. 2.9), which would otherwise suggest that their average temperatures are not significantly biased. Although the statistical uncertainties are large, the Lieu et al. (2016) $M_{500} - T_{300\mathrm{kpc}}$ relation has a significantly higher normalisation than the $L_{500} - T_{500}$ relations based on X-ray hydrostatic masses. The difference in aperture may have a small effect at high temperatures, since a 300 pkpc aperture is smaller than r_{500} for the majority of $T > 1$ keV systems and could yield a slightly higher temperature measurement due to the declining temperature profiles of such systems. However, this would lower rather than boost the normalisation of the $M_{500} - T$ relation. Furthermore, the difference between these measures is small in our simulations ($\lesssim 0.1$ dex) and Giles et al. (2016) find no systematic differences between the two temperatures for the XXL-100-GC sample. This implies that the offset between the X-ray and weak lensing calibrated $M_{500} - T$ relations is due to an offset in halo mass. Under the assumption that weak lensing masses are unbiased, the Lieu et al. (2016) results suggest that X-ray hydrostatic masses are biased low and our agreement with their results suggests that FABLE systems possess realistic global temperatures.

We find that the simulation predictions of the $M_{500} - T$ relation are rather robust with respect to variations of the feedback modelling. The better agreement of the simulations with the weak lensing calibrated relation may hence provide further circumstantial evidence that weak lensing mass measurements are less biased. We caution, however, that changes in the feedback modelling beyond those considered here may result in larger variations in the predicted normalisation.

2.5.6 *SZ-Mass relation*

In Fig. 2.10 we plot the mean SZ flux–halo mass relation calculated as described in Section 2.3.4. As in Planck Collaboration XI (2013), we self-similarly scale the SZ flux to redshift $z = 0$ and to a fixed angular diameter distance of 500 Mpc, yielding the SZ signal, $\tilde{\mathrm{Y}}_{5\mathrm{r}_{500}}$, in units of square arcminutes.

The FABLE simulations produce a power-law relation extending from massive galaxies to clusters in good agreement with both the original Planck Collaboration XI (2013) relation, which is based on halo masses derived from a semi-analytic galaxy formation model, and the weak lensing calibrated relation from Wang et al. (2016). At $\sim 5 \times 10^{12} M_{\odot}$ there is a slight upturn in the observed relation not

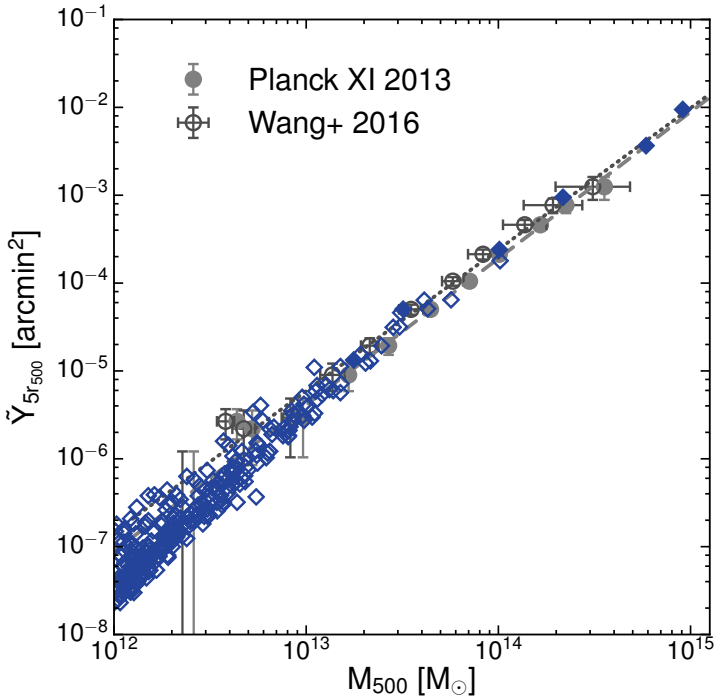


Figure 2.10: SZ flux as a function of halo mass at $z = 0$. The SZ flux is calculated as described in Section 2.3.4 and has been scaled to $z = 0$ and a fixed angular diameter distance of 500 Mpc. Symbols with error bars show the mean SZ signal in bins of halo mass for a sample of SDSS locally brightest galaxies presented in Planck Collaboration XI (2013). Filled circles correspond to halo masses derived by Planck Collaboration XI (2013) from a semi-analytic galaxy formation model and open circles show the recalibrated halo masses and associated uncertainties determined by Wang et al. (2016) using stacked weak lensing analyses. The dashed line shows the best-fitting relation from Planck Collaboration XI (2013) and the dotted line shows the best-fitting relation for the Wang et al. (2016) recalibration.

seen in the simulations, however, these two mass bins correspond to detections of less than two sigma. Furthermore, Planck Collaboration XI (2013) state that the three lowest mass bins are noticeably affected by dust contamination and may be more uncertain than the statistical measures indicate. Indeed, Greco et al. (2015) explicitly model the dust emission from each LBG in the sample and find that, for the low-mass LBGs with $M_{500} \lesssim 10^{13.3} M_\odot$, the stacked signal from dust emission is comparable to or larger than the stacked SZ signal.

2.6 ICM PROFILES

2.6.1 *Density profiles*

In Fig. 2.11 we show spherically-averaged radial density profiles of the ICM in FABLE groups and clusters at $z = 0$. In the left hand panel we compare group-scale

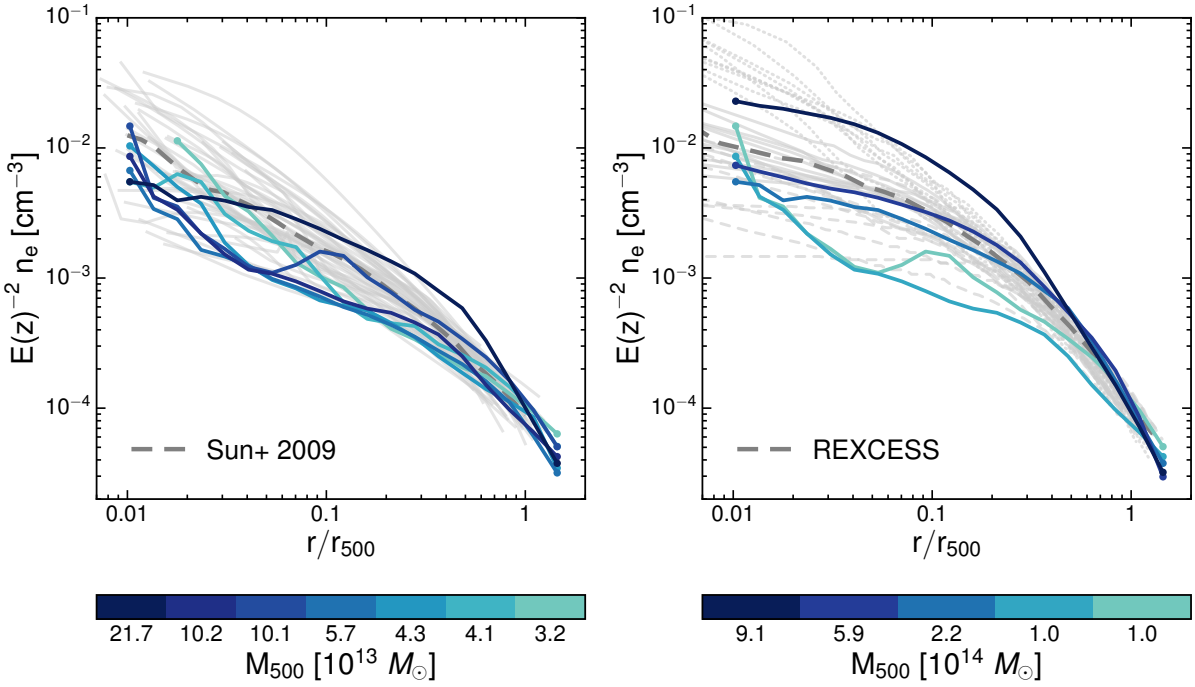


Figure 2.11: Density profiles of the ICM in FABLE groups and clusters at $z = 0$ in comparison to observed profiles. Lines show individual profiles of simulated systems colour coded by halo mass. All profiles have been self-similarly scaled in redshift. In the left hand panel we plot the density profiles of the Sun et al. (2009) galaxy groups (grey lines) and the profiles of a mass-selected sample of simulated group-scale systems with a similar median halo mass (see main text). The thick dashed line shows the median of the Sun et al. (2009) sample. In the right hand panel we show density profiles for all cluster-scale systems with $\mathrm{M}_{500} \geq 10^{14} M_\odot$ in comparison to those of the REXCESS cluster sample (grey lines; Croston et al. 2008). For the observed sample, solid lines correspond to relaxed clusters, dashed lines to disturbed clusters and dotted lines to cool-core clusters according to the definitions of Pratt et al. (2009). The thick dashed line shows the median REXCESS profile.

systems to the density profiles of the Sun et al. (2009) sample of groups observed with *Chandra*. The Sun et al. (2009) sample consists of 43 groups with X-ray hydrostatic mass estimates in the range $1.6 \times 10^{13}\ M_\odot \leq M_{500} \leq 1.8 \times 10^{14}\ M_\odot$ with a median mass of $7.3 \times 10^{13} M_\odot$. The sample was drawn from *Chandra* archival data with the requirement that gas properties could be derived out to at least r_{2500} ($\approx 0.47\ r_{500}$). We compare to a halo mass-selected sample of simulated groups with $3.2 \times 10^{13}\ M_\odot \leq M_{500} \leq 2.2 \times 10^{14}\ M_\odot$ and a median mass of $5.7 \times 10^{13} M_\odot$. The simulated density profiles are in good agreement with the Sun et al. (2009) profiles beyond $\sim 0.3\, r_{500}$ but slightly underestimate the median density at smaller radii. At $\gtrsim 0.3\, r_{500}$ the agreement is consistent with our match to observed gas mass fractions (see Fig. 2.5), as this region contains ~ 90 per cent of the total gas mass.

In Section 2.5 we argued that if there is a significant X-ray mass bias then the FABLE systems may be over-luminous in X-rays and that this may be the result of an overabundance of gas (the observed gas fractions in Fig. 2.5 being biased high). In this case, we would expect to overestimate the gas density in the outer regions compared to the data. The fact that this is not seen does not, however, rule out the possibility of a significant X-ray mass bias. One explanation is the difference between the halo mass distributions of the observed and simulated samples. Although we have tried to match the median halo masses of the observed and simulated samples as closely as possible given the small sample size, the median mass of the simulated sample is 22 per cent lower than that of the Sun et al. (2009) sample and would decrease further in the case of an X-ray mass bias (a 45 per cent difference if X-ray masses are biased low by 30 per cent). As the gas content of massive haloes is a relatively strong function of their halo mass (see e.g. Fig. 2.5), this could result in a mismatch between the two samples. Another explanation is that the average X-ray luminosity of the Sun et al. (2009) sample is biased high due to their requirement that group emission be traced out to a significant fraction of r_{500}. Indeed, as we discussed in Section 2.5.4, the Sun et al. (2009) groups with $T_{500} \sim 1$ keV possess slightly higher X-ray luminosities compared to the Osmond & Ponman (2004) and Zou et al. (2016) samples, such that the $L^{bol}_{500} - T_{500}$ relation of our simulated groups is actually in good agreement with the Sun et al. (2009) data. This could explain why our simulated groups match the density profile of the Sun et al. (2009) groups yet seem to overestimate the X-ray luminosity relative to other studies.

At $r \lesssim 0.3\, r_{500}$ the simulated density profiles lie within the observed scatter but largely fall below the median observed profile. This suggests that our AGN feedback model may displace slightly too much gas from the central regions. On the other hand, the observed groups are detected via their X-ray flux, which may preferentially select systems with high central densities compared to our halo mass-selected sample, particularly if the Sun et al. (2009) sample is biased toward high luminosities relative to other X-ray selected samples.

In the right hand panel of Fig. 2.11 we compare our simulated clusters to the density profiles of the REXCESS clusters, a representative sample of 31 clusters observed with *XMM-Newton* (Böhringer et al., 2007; Croston et al., 2008). The REXCESS clusters were chosen such that r_{500} lies well within the field of view of *XMM-Newton*, allowing detailed local background modelling and increased measurement precision at large radii (Croston et al., 2008). The sample is unbiased with respect to cluster morphology or dynamical state, containing a representative distribution of relaxed, cool-core and morphologically disturbed clusters (as

defined in Pratt et al. 2009), which correspond to the solid, dotted and dashed lines in Fig. 2.11, respectively. The REXCESS sample covers the mass range $1.0 \times 10^{14} M_{\odot} \leq M_{500} \leq 7.8 \times 10^{14} M_{\odot}$ with a median mass of $2.6 \times 10^{14} M_{\odot}$. Our comparison sample consists of all five FABLE clusters with $M_{500} \geq 1.0 \times 10^{14} M_{\odot}$ and has a comparable median mass of $2.1 \times 10^{14} M_{\odot}$.

Overall the density profiles of our simulated clusters are a good match to the REXCESS clusters. At $r \gtrsim 0.3\, r_{500}$ the densities of the three most massive FABLE clusters are somewhat high compared to the median observed profile, which is consistent with the excess gas we might expect in our simulated clusters in the case of an X-ray mass bias. Indeed, the cumulative gas fraction (not shown) rises more steeply between ~ 0.1–$0.6\ r_{500}$ than the REXCESS profiles (Pratt et al., 2010). This suggests that AGN feedback may act too violently, pushing gas from the cluster centre and causing it to pile up at larger radii. A similar trend was found for clusters in IllustrisTNG (Barnes et al., 2018b).

The simulated profiles are similar in shape to the relaxed or disturbed REXCESS clusters. Of the five simulated clusters only one has a central density comparable to observed (weak) cool-core clusters. Potentially, heating of the cluster centre by AGN feedback might be preventing cool-cores from forming in the same proportion as observed at $z \sim 0$ (10 of 31 REXCESS clusters and ~ 30–40 per cent in SZ surveys, e.g. Planck Collaboration XI 2011; Andrade-Santos et al. 2017; Rossetti et al. 2017). A larger sample will be needed to assess this in detail. Reproducing the observed fraction of cool-core galaxy clusters in cosmological simulations with feedback is a notoriously difficult problem (e.g. Borgani & Kravtsov 2011; Kravtsov & Borgani 2012). There has been some recent success in this area (e.g. Rasia et al. 2015; Barnes et al. 2018b; Hahn et al. 2017), however, these simulations tend to underestimate the observed cool-core fraction when compared to low-redshift SZ-selected samples. It is clear that an AGN feedback model that is capable of reproducing the global properties of clusters does not necessarily provide an effective description of the physical processes responsible for the creation of cool cores: additional processes such as AGN-driven turbulence, cosmic rays, stabilisation from magnetic fields or anisotropic thermal conduction may be required.

2.6.2 *Temperature profiles*

In Fig. 2.12 we plot the dimensionless temperature profiles at $z = 0$ for the same group- and cluster-scale samples described in the previous section. We facilitate a comparison between different halo masses by normalising the profiles

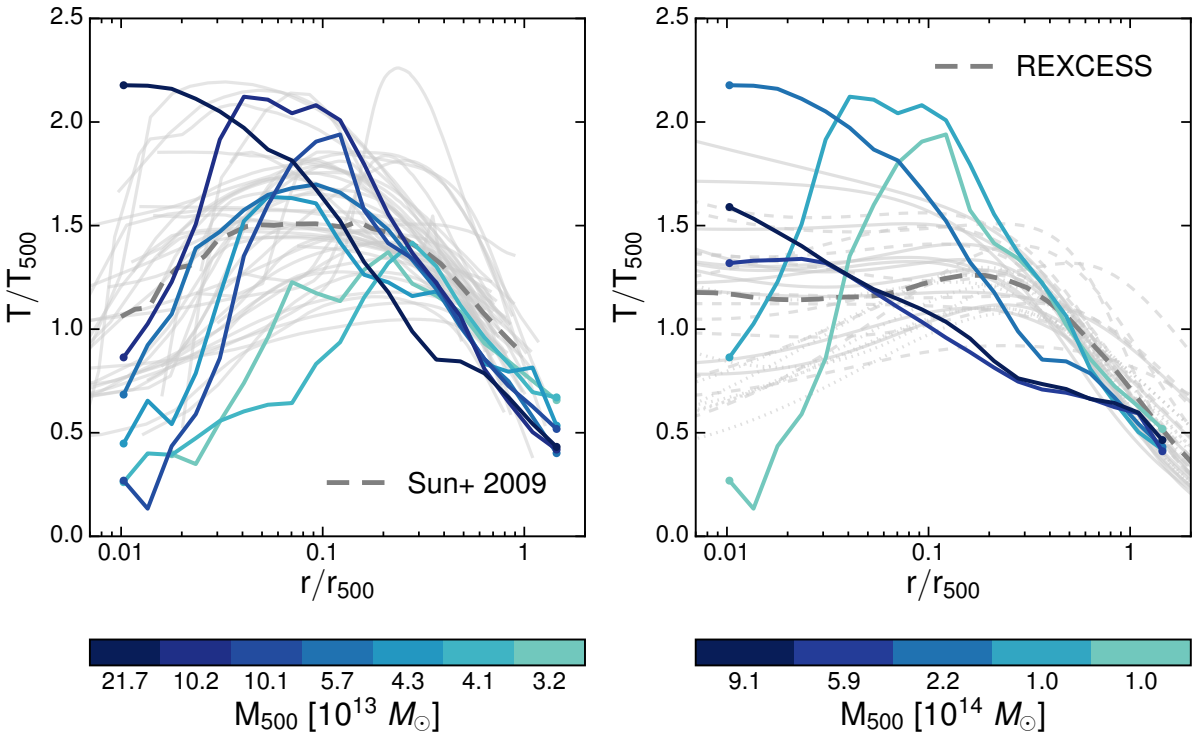


Figure 2.12: Dimensionless temperature profiles at $z = 0$ for the same simulated and observed samples as shown in Fig. 2.11. Profiles are normalised by the characteristic temperature defined in equation 2.2. In the left hand panel, solid grey lines show the individual deprojected temperature profiles of the Sun et al. (2009) sample. The thick dashed line shows the median profile. In the right hand panel, grey lines show individual temperature profiles for REXCESS clusters derived from the best-fitting pressure (Arnaud et al., 2010) and entropy (Pratt et al., 2010) profiles. Line styles are the same as shown in Fig. 2.11.

by the characteristic temperature,

$$kT_{500} = \mu m_p G M_{500}/2r_{500}, \tag{2.2}$$

the temperature of an isothermal sphere of mass M_{500} and radius r_{500}. Here μ is the mean molecular weight, which we take as $\mu = 0.59$, and m_p is the proton mass.

In the left hand panel of Fig. 2.12 we compare to the deprojected temperature profiles of the Sun et al. (2009) groups. We find that beyond the core ($\gtrsim 0.2\,r_{500}$) the temperature profiles of the FABLE groups have a similar slope to the median observed profile. The normalisation is somewhat lower than observed, however, this may be the result of an X-ray hydrostatic mass bias in the Sun et al. (2009) mass estimates: since the halo mass, M_{500}, is related to the characteristic temperature scaling as $T_{500} \propto M_{500}^{2/3}$, a bias toward lower masses would shift the observed dimensionless temperature profiles to higher values. Within the core, the simulated

temperature profiles show similar scatter to observed and in general follow the same shape as the observed profiles.

For the most massive simulated system in the group sample, the temperature rises steadily towards the centre rather than dropping within the core. This may be a side-effect of our relatively simple model for radio-mode AGN feedback, which injects bubbles of thermal energy at irregular intervals. In reality, such bubbles are expected to be supported by non-thermal pressure from, e.g., cosmic rays and should only contribute to the observed temperature profile once the injected energy has thermalised.

In the right hand panel of Fig. 2.12 we compare to the dimensionless temperature profiles of the REXCESS clusters. The REXCESS temperature profiles rise with roughly constant slope from r_{500} down to $\sim 0.3\,r_{500}$ before dropping slowly or levelling out within the cluster core. Close to r_{500} there is a slight offset between the predicted and observed temperatures. As for the Sun et al. (2009) group comparison, this could be explained by a bias in the X-ray hydrostatic mass estimates. Between ~ 0.3–$1\,r_{500}$, three of the five simulated temperature profiles show a similar slope to observed but for the two most massive clusters in the sample the slope is somewhat shallower.

The REXCESS temperature profiles are not as strongly peaked as the Sun et al. (2009) profiles. Hence the two FABLE systems with $M_{500} \sim 1 \times 10^{14} M_{\odot}$, which are included in both the group- and cluster-scale samples, are a reasonable match to one or more of the Sun et al. (2009) systems but reach higher peak temperatures compared to the REXCESS clusters. This is unsurprising given that the halo masses of these systems are on the boundary between groups and clusters and correspond only to the very lowest masses in the REXCESS sample. The three most massive FABLE clusters possess temperature profiles that continue to rise within the cluster core ($\lesssim 0.1\,r_{500}$), unlike the observed temperature profiles, which tend to level out. As noted above, this could be because AGN bubble feedback is modelled via injection of thermal energy rather than of non-thermal components. This may result in an overheating of the ICM in the central regions.

2.6.3 *Entropy profiles*

In Fig. 2.13 we plot the dimensionless entropy profiles at $z = 0$ for the same group- and cluster-scale samples described in Section 2.6.1. We use the widely adopted definition of ICM "entropy", $K = kT/n_e^{2/3}$, and normalise by the characteristic entropy scale, $K_{500} = kT_{500}/n_{e,500}^{2/3}$, which reflects the mass variation expected in a self-similar model. The characteristic temperature kT_{500} is defined

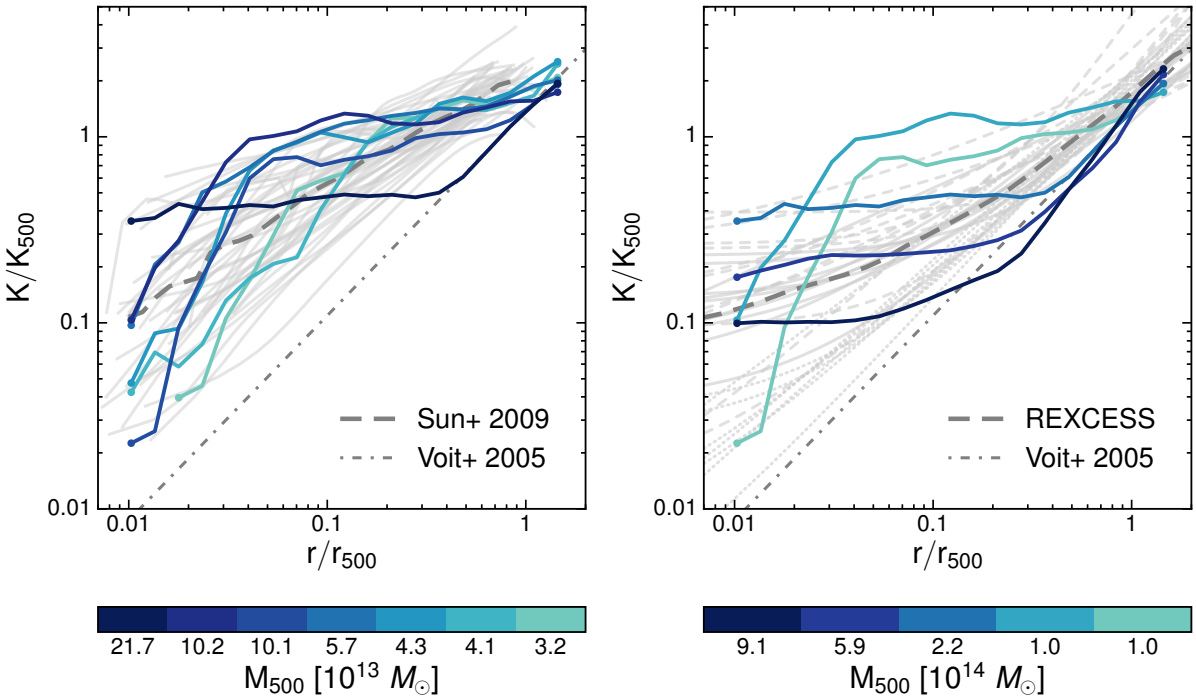


Figure 2.13: Dimensionless entropy profiles at $z = 0$ for the same simulated and observed samples as shown in Fig. 2.11. Profiles are scaled by the characteristic entropy scale K_{500} as defined in the text. The dash-dotted line shows the baseline ICM entropy profile derived by Voit (2005) from non-radiative simulations. In the left hand panel, grey lines show the entropy profiles of the Sun et al. (2009) groups derived from the density and temperature profiles shown in Fig. 2.11 and 2.12. In the right hand panel, grey lines show the best-fitting entropy profiles of the REXCESS clusters (Pratt et al., 2010). Line styles are the same as shown in Fig. 2.11.

in equation 2.2. $n_{e,500}$ is the mean electron density within r_{500} and is defined as

$$n_{e,500} = 500 f_b \rho_c(z)/(\mu_e m_p), \tag{2.3}$$

assuming the global baryon fraction $f_b = \Omega_b/\Omega_{\rm M}$ in a universe with critical density $\rho_c(z)$ at redshift z and a mean molecular weight per free electron μ_e. We adopt a value of $\mu_e = 1.14$ and use f_b and $\rho_c(z)$ corresponding to our assumed cosmology.

In the left hand panel of Fig. 2.13 we plot the entropy profiles of the group-scale systems. Close to r_{500} the entropy profiles tend to the power-law prediction from Voit (2005), which was derived from hydrodynamic simulations in the absence of non-gravitational processes. At smaller radii the entropy profiles deviate from this relation due to non-gravitational processes such as AGN feedback, which ejects and heats gas in the group centre, and star formation, which removes low entropy gas. The simulated entropy profiles are in good agreement with the Sun et al. (2009) profiles over the full range of radii. The exception is the most massive system, which follows the non-radiative relation before quickly flattening into an

isentropic core at $\sim 0.4\,r_{500}$. This is consistent with the temperature profile of this system, which rises steadily towards the cluster centre instead of dropping slowly within the core.

In the right hand panel of Fig. 2.13 we see that the observed entropy profiles of the REXCESS clusters run approximately parallel to the Voit (2005) relation at large radii and slowly flatten toward the cluster centre. This change in slope occurs less rapidly than in the less massive Sun et al. (2009) groups, since the deeper potentials of more massive haloes make non-gravitational processes such as AGN feedback less effective. The cool-core REXCESS clusters have mostly power-law-like entropy profiles while the disturbed clusters tend to have cored entropy profiles.

The two FABLE clusters with $M_{500} \sim 1 \times 10^{14} M_{\odot}$ have fairly flat entropy profiles down to $\sim 0.05\,r_{500}$ before dropping rapidly within the core. These profiles lie within the intrinsic scatter of the Sun et al. (2009) sample but are unlike the entropy profiles of the REXCESS clusters. This suggests that these systems are more similar to high-mass groups than low-mass clusters. For the three most massive FABLE clusters, the entropy at $\gtrsim 0.3\,r_{500}$ closely follows the baseline relation of Voit (2005). This suggests that AGN feedback has little effect on the thermodynamics of the gas at large radii. The observed profiles lie somewhat higher than the baseline and the simulations, however, the observed profiles may be slightly overestimated in the case of an X-ray mass bias. At smaller radii the three most massive clusters show flat central entropy distributions. These cored entropy profiles are characteristic of many of the relaxed and disturbed REXCESS clusters, although the cores extend to somewhat larger radii than observed, which is likely related to the slight overprediction of the gas density in these systems at $\sim 0.3\,r_{500}$ (see Sec. 2.6.1). This is consistent with a picture in which AGN feedback in FABLE clusters is overly effective at heating and expelling gas in the central regions but is relatively ineffective at large radii.

2.6.4 *Pressure profiles*

In Fig. 2.14 we plot the pressure profiles of the ICM at $z = 0$ for the group- and cluster-scale samples described in Section 2.6.1. We normalise the pressure profiles, $P(r) = kT(r)n_e(r)$, by the characteristic pressure, $P_{500} = kT_{500}n_{e,500}$, where kT_{500} and $n_{e,500}$ are defined in equations 2.2 and 2.3.

In the left hand panel of Fig. 2.14 we compare to the Sun et al. (2009) pressure profiles. At $r \gtrsim 0.3\,r_{500}$ the simulated profiles are a good match to the data. Similar to the dimensionless temperature and entropy profiles, there is a slight offset in

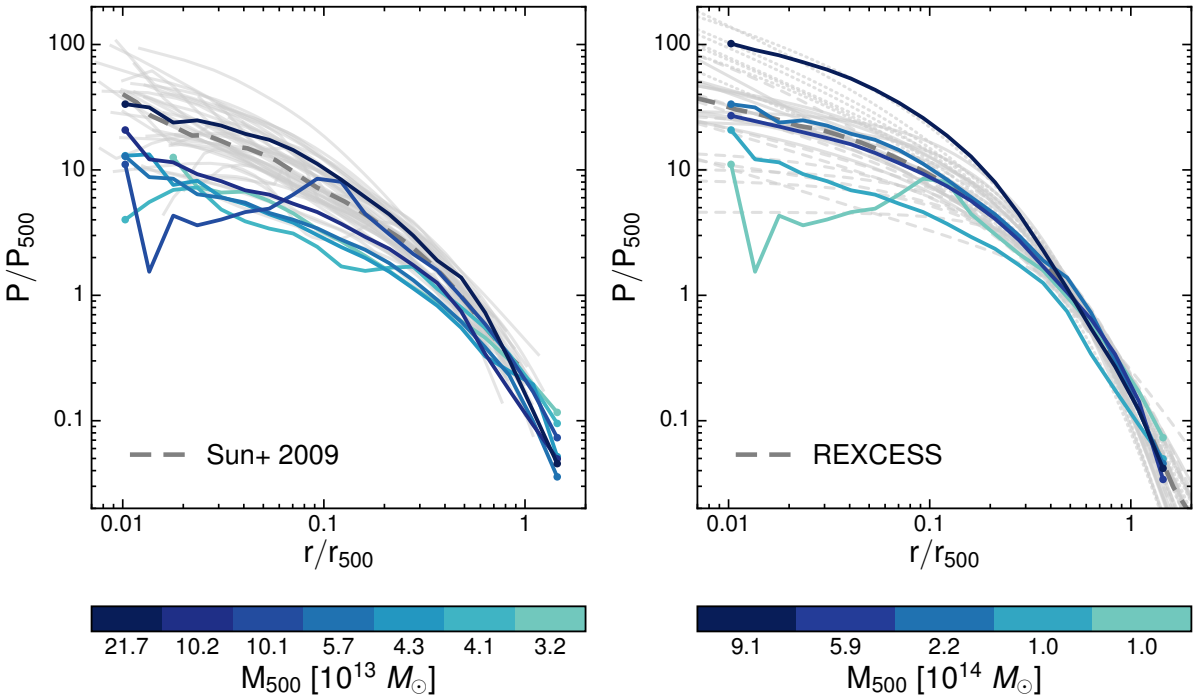


Figure 2.14: Dimensionless pressure profiles at $z = 0$ for the same simulated and observed samples as shown in Fig. 2.11. Profiles are scaled by the characteristic pressure P_{500} as defined in the text. In the left hand panel, grey lines show the pressure profiles of the Sun et al. (2009) groups derived from the density and temperature profiles shown in Fig. 2.11 and 2.12. In the right hand panel, grey lines show the best-fitting pressure profiles of the REXCESS clusters (Pratt et al., 2010). Line styles are the same as shown in Fig. 2.11.

normalisation with respect to the median observed profile, possibly due to an X-ray mass bias. Inside $\sim 0.3\, r_{500}$ the most massive system remains in good agreement with the data, however the central pressure in the less massive simulated groups is slightly underestimated. This is consistent with the density profiles shown in Fig. 2.11, which are slightly underestimated at $r \lesssim 0.3\, r_{500}$ compared to the median observed profile. As we discussed in Section 2.6.1, this suggests that AGN feedback may be ejecting too much gas from the central regions of galaxy groups, although selection effects in the Sun et al. (2009) sample may also play a role.

In the right hand panel of Fig. 2.14 we find excellent agreement with the REXCESS pressure profiles across the full range of radii. In the outskirts of the FABLE clusters ($r \geq 0.5\, r_{500}$) the dimensionless pressure profiles coincide over a wide range of halo masses. The same is true for the observed clusters, which suggests that both the simulated and observational samples represent fairly self-similar populations. At small radii there is a departure from the self-similar scaling due to the effects of non-gravitational processes. The dispersion increases toward the cluster centre in both the observed and simulated samples to a similar degree.

2.7 DISCUSSION

The FABLE simulations employ an updated version of the Illustris galaxy formation model. Specifically we have updated the sub-grid models for feedback from stars and AGN in order to reproduce the $z = 0$ galaxy stellar mass function and the present-day gas mass fractions of massive haloes. The latter were not considered during the Illustris calibration and were severely underestimated with respect to observations. By adopting a model that reproduces observed gas fractions, we have obtained significantly more realistic galaxy groups and clusters while, at the same time, maintaining a good match to the observed galaxy stellar mass function in the field.

The FABLE model produces a very similar galaxy stellar mass function to Illustris (see Fig. 2.2), despite significantly changing the way in which AGN feedback regulates star formation in massive galaxies. In Illustris, quasar-mode feedback was continuous such that a relatively small amount of accreted feedback energy was injected into the surrounding gas at every timestep. This often resulted in too little thermal energy being input into too much gas, allowing the energy to be efficiently radiated away. The quasar-mode was therefore inefficient at suppressing star formation and strong radio-mode feedback was needed to suppress stellar mass buildup in massive haloes. This was achieved with a long duty cycle for the radio-mode such that feedback events were infrequent but highly energetic. The major side-effect of this model was that the radio-mode acted too violently on the gas, resulting in severely underestimated gas fractions in Illustris at $z = 0$ (see Fig. 2.5). In FABLE we have introduced a modification to the quasar-mode in the form of a duty cycle. Rather than a continuous injection of thermal energy as in Illustris, the available feedback energy is stored over a 25 Myr time period before being injected into the surrounding gas in a single event. This heats the gas to much higher temperatures, resulting in a longer cooling time and overall more efficient feedback. The updated quasar-mode feedback is more effective at suppressing the stellar mass buildup of massive galaxies, thereby allowing the radio-mode to operate on a much shorter duty cycle. This gentler form of radio-mode feedback has a smaller impact on the gas content of massive haloes and has enabled us to produce groups and clusters with realistic gas fractions (see Fig. 2.5). We point out that the agreement with observed gas fractions at cluster scales ($M_{500} \gtrsim 10^{14} M_{\odot}$) was not guaranteed, since the feedback model was not calibrated at such scales. By applying our calibrated model to cluster-scale objects simulated using the zoom-in technique, we are able to compare the predictions of the FABLE model to a variety of observational constraints across a wide range

of halo masses. We demonstrate very good agreement with observations for a number of group and cluster properties, including stellar mass fractions, X-ray luminosity–mass relations, integrated Sunyaev–Zel'dovich flux and radial profiles of the intracluster medium.

Yet there remain some discrepancies with our model compared to observations. In particular we find that the X-ray luminosity–temperature ($L_{500} - T$) relation lies on the upper end of the observed scatter (see Fig. 2.8). From the $L_{500} - T$ relation alone it is unclear whether the cause of this offset is dominated by overestimated X-ray luminosities or underestimated spectroscopic temperatures. We aim to gain some insight on the discrepancy by comparing to other observed scaling relations such as the halo mass–temperature ($M_{500} - T$) relation and the X-ray luminosity–halo mass ($L_{500} - M_{500}$) relation. However, the difference between relations based on X-ray hydrostatic mass estimates versus weak lensing mass estimates means that the conclusion is dependent upon which is used for the comparison.

If we compare the $M_{500} - T_{500}$ relation of the FABLE simulations to observational data based on X-ray derived halo masses (left hand panel of Fig. 2.9), we would conclude that the average X-ray temperatures of our systems are systematically underestimated by ~ 0.2 dex. This is large enough to explain the discrepancy in $L_{500} - T_{500}$ without affecting the $L_{500} - M_{500}$ relations, which are in excellent agreement with observed $L_{500} - M_{500}$ relations based on X-ray derived halo masses (Fig. 2.6 and 2.7).

On the other hand, we have very good agreement with the $M_{500} - T_{300\mathrm{kpc}}$ relation of Lieu et al. (2016), which uses halo masses measured via weak lensing (right hand panel of Fig. 2.9). This would suggest that our simulated systems possess realistic global temperatures. Similarly, our agreement with the weak lensing calibrated $Y - M_{500}$ relation (Fig. 2.10) implies that the mass-weighted temperatures of our simulated systems are realistic (the integrated SZ flux being proportional to the total thermal energy content of the gas). We have confirmed that our spectroscopic temperature estimates are not systematically lower than the mass-weighted temperature for the mass range in question, which provides further evidence that the discrepancy in $L_{500} - T_{500}$ relation is unlikely to be due to underestimated temperatures.

There is a significant offset in normalisation between the weak lensing $M_{500} - T$ relation of Lieu et al. (2016) and the $M_{500} - T$ relations based on X-ray masses, albeit with large scatter in the weak lensing data. In fact, Lieu et al. (2016) perform a comparison of the normalisation of different $M_{500} - T$ relations from the literature and find that relations based on weak lensing masses favour ~ 40 per cent higher normalisations than those based on X-ray hydrostatic masses. This implies a

systematic difference between halo masses measured from weak lensing and masses measured from X-rays. Indeed, a number of observational studies have found that, within r_{500}, X-ray hydrostatic masses are biased low compared to weak lensing masses by ~ 25–30 per cent (e.g. Donahue et al. 2014; von der Linden et al. 2014b; Hoekstra et al. 2015; Simet et al. 2017). A slightly larger X-ray mass bias of ~ 40 per cent, as seemingly preferred by the Lieu et al. (2016) sample, is large enough to reconcile the results of cluster abundance studies with cosmological constraints from *Planck* measurements of the primary CMB (Planck Collaboration XXIV, 2016).

Under the assumption that weak lensing masses are less biased than X-ray hydrostatic masses, we would deduce that the spectroscopic temperatures of FABLE groups and clusters are realistic and that the discrepancy in $L_{500} - T_{500}$ is largely the result of overestimated X-ray luminosities. This is consistent with the $L_{500} - M_{500}$ relations shown in Fig. 2.6 and 2.7, which suggest that the predicted X-ray luminosities may be overestimated as a function of halo mass compared to observations based on weak lensing (rather than X-ray) mass measurements. Furthermore, if X-ray hydrostatic masses are biased low, then the observational constraints on stellar and gas mass fractions plotted in Fig. 2.4 and 2.5 would be biased high. For example, Eckert et al. (2016) find that the weak lensing calibrated gas fraction of XXL-100-GC clusters is significantly lower than independent results based on X-ray hydrostatic masses. Since we have calibrated our feedback model to reproduce observed gas mass fractions assuming a negligible X-ray mass bias, this would imply that FABLE groups and clusters are too gas rich.[3] The excess gas could then explain our overpredicted X-ray luminosities at fixed temperature.

On the other hand, a large X-ray hydrostatic mass bias ($\gtrsim 30$ per cent) implies a baryon depletion factor significantly exceeding that predicted by numerical simulations (e.g. Eckert et al. 2016). Moreover, a number of studies find results consistent with little to no bias (e.g. Gruen et al. 2014; Israel et al. 2014; Smith et al. 2016; Applegate et al. 2016; Maughan et al. 2016; Andreon et al. 2017).

Few cosmological hydrodynamical simulations manage to convincingly reproduce the observed X-ray scaling relations. The cosmo-OWLS and BAHAMAS projects obtain a good match to the observed X-ray luminosity–halo mass relation with relatively low-resolution simulations and when modelling a significant X-ray mass bias. At much higher resolution, the C-EAGLE clusters, which employ the

[3]We caution that if M_{500} as measured from X-rays is biased low then the corresponding estimate of r_{500}, and by extension the gas masses and other quantities measured within this radius, will also be biased low. This should be taken into consideration when comparing the simulation results to observations based on potentially biased X-ray masses.

EAGLE galaxy formation model, are slightly over-luminous for a given halo mass due to the clusters being too gas rich. Similarly, the hydrodynamical cluster simulations presented in Truong et al. (2018) possess lower than observed temperatures and approximately 30 per cent higher X-ray luminosities than observed, although the latter discrepancy they suggest is at least partly due to sample selection. Mesh-based cosmological hydrodynamical simulations encounter similar issues. For example, the Rhapsody-G (Hahn et al., 2017) suite of cluster zoom-in simulations, which use an Eulerian adapative mesh refinement (AMR) method, show X-ray luminosities as a function of halo mass consistently higher than observed (by ~ 20 per cent at $M_{500} \approx 10^{15} M_{\odot}$ and about a factor of 2 at $M_{500} \approx 10^{14} M_{\odot}$). Interestingly, Hahn et al. (2017) find that the normalisation of the X-ray luminosity–halo mass relation is insensitive to the AGN feedback parameters, including drastic changes to the length of the duty cycle, contrary to the results of SPH simulations with a similar AGN feedback model (e.g. Le Brun et al. 2014).

In our simulations, which are run with the AREPO moving-mesh code, we find that changes to the duty cycle and energetics of AGN feedback can have a large impact on the gas mass fractions of massive haloes (see Appendix A). We expect that the FABLE model could likely be adjusted to produce somewhat lower gas mass fractions and thus lower X-ray luminosities in massive haloes by lengthening the duty cycle of radio-mode feedback. This would make individual events more energetic and therefore more effective at ejecting gas beyond the virial radius. In Appendix A we show that increasing the burstiness of the radio-mode in this way can significantly lower halo gas fractions without drastically altering the $z = 0$ galaxy stellar mass function, which is already in good agreement with observations in our fiducial model.

On the other hand, further tuning of our relatively simplistic model for thermal bubble feedback is unlikely to significantly improve the ICM profiles of our simulated clusters, which show indications of over-heating and removal of too much gas in the central regions, while gas at $\gtrsim 0.5\, r_{500}$ is relatively unaffected (see e.g. entropy profiles in Section 2.6.3). A similar though more extreme predicament applies to the C-EAGLE clusters (Barnes et al., 2017b), which show lower than observed gas density in the core and entropy profiles with significantly larger cores and higher central entropies than observed. Barnes et al. (2017b) suggest that AGN feedback in C-EAGLE is too active at late times, increasing the central entropy of clusters and preventing the formation of cool-core systems with steep central density and entropy profiles. The FABLE clusters also do not contain an obvious strong cool-core system, although a larger sample will be needed to assess this issue in detail. Other numerical works such as Rasia et al. (2015) and Hahn

et al. (2017) have reproduced the observed dichotomy between cool-core and non-cool-core clusters, however, this does not necessarily imply good agreement with observational constraints on the global properties of clusters, since both models tend to produce clusters with somewhat higher than observed X-ray luminosities (Hahn et al., 2017; Truong et al., 2018). IllustrisTNG also produce a fraction of cool-core clusters that is in agreement with observations between $0.25 < z < 1.0$, however, the cool-core fraction is underpredicted at $z < 0.25$, with more clusters showing close to isentropic cores at $z = 0$ than observed (Barnes et al., 2018b).

Given that almost all numerical simulations are unable to convincingly reproduce the observed thermodynamic profiles of cluster core regions, it seems vital that simulation models for AGN feedback should continue to be improved, for example, by inflating bubbles self-consistently with AGN jet feedback (Bourne & Sijacki, 2017; Weinberger et al., 2017a), while also incorporating additional physical processes that have previously been neglected, such as cosmic-rays (Jacob & Pfrommer, 2017; Pfrommer et al., 2017), outflows driven by radiation pressure from AGN (Costa et al., 2018b,a; Ishibashi et al., 2018) and anisotropic thermal conduction (Kannan et al., 2016, 2017). In addition, further improvement to the realism of simulated groups and clusters will rely on a better understanding of the issues of mass bias and selection effects so that reliable comparisons to unbiased observational data sets can be performed.

2.8 CONCLUSIONS

In this chapter we have introduced the FABLE suite of cosmological hydrodynamical simulations, which consists of a $(40\ h^{-1}\ \mathrm{Mpc})^3$ cosmological volume and a number of zoom-in simulations of individual galaxy groups and clusters. The simulations were performed with the moving mesh code AREPO and an updated version of the Illustris galaxy formation model. We have adapted the sub-grid models for stellar and AGN feedback in order to reproduce galaxy groups and clusters with more realistic gas fractions compared to Illustris whilst maintaining a similarly high level of agreement with the observed present-day galaxy stellar mass function. In this chapter we have presented various other comparisons with observations, including X-ray and SZ scaling relations and radial profiles of the ICM over a wide range of halo masses. Our main conclusions are as follows:

- We obtain very good agreement with observed galaxy stellar mass functions. The high mass end of the $z = 0$ mass function is similar to that of Illustris, despite significant changes to the way in which AGN feedback suppresses

the buildup of stellar mass. While the FABLE model was calibrated to reproduce the $z \approx 0$ mass function, the fact that the agreement with observations continues to higher redshift is a success of the model.

- The stellar mass fractions of FABLE galaxy groups and clusters are also an excellent match to low-redshift observations, including in massive clusters that were not present in the calibration volume.

- The $z = 0$ halo gas mass fractions represent a major improvement over Illustris and are now in good agreement with observations. This can be attributed to much less energetic but more frequent thermal energy injections in the radio-mode of AGN feedback, which remove less gas from haloes compared to the Illustris model.

- The predicted X-ray luminosity–total mass ($L_{500} - M_{500}$) relations are in excellent agreement with observed relations based on X-ray hydrostatic mass estimates but seem to overestimate the X-ray luminosity for a given halo mass when compared with weak lensing mass estimates. The difference between observed relations is consistent with a significant X-ray mass bias. Similarly, a comparison of observed total mass–spectroscopic temperature ($M_{500} - T$) relations reveals a systematic difference between those based on X-ray hydrostatic masses and weak lensing masses. The FABLE simulations are in good agreement with $M_{500} - T$ data based on weak lensing masses but have significantly lower global temperatures/higher masses compared to relations using only X-ray data.

- The slope of the predicted X-ray luminosity–spectroscopic temperature ($L_{500} - T_{500}$) relation is in excellent agreement with observations. The normalisation of the relation lies, however, on the upper end of the scatter in the data. The size of this offset is similar to the offset with weak lensing-based $L_{500} - M_{500}$ relations and X-ray-only $M_{500} - T_{500}$ relations. This implies that the discrepancy in $L_{500} - T_{500}$ could be due to either overestimated X-ray luminosities or underestimated global temperatures. We lean towards the former explanation, as this is consistent with the general expectation that weak lensing masses are less biased than X-ray hydrostatic masses. An improved understanding of mass bias will be important for making further progress here.

- The simulations are also in excellent agreement with the mean Sunyaev–Zel'dovich flux–total mass ($Y_{5r500} - M_{500}$) relation derived from *Planck* obser-

vations of locally bright galaxies. This implies that the global temperatures of our simulated systems are not significantly underestimated, consistent with our match to the weak lensing $M_{500} - T$ relation.

- In general, the radial profiles of the ICM are a good match to observations outside $\sim 0.3\, r_{500}$, where the majority of the ICM is located. Density and pressure profiles of the ICM are in good agreement with observations of both group- and cluster-scale systems. The group-scale profiles have slightly lower-than-observed density/pressure within $\sim 0.3\, r_{500}$, however, this may be (partly) due to selection effects in the observed sample. The temperature and entropy profiles of $\lesssim 10^{14} M_{\odot}$ haloes are also in good agreement with observations, while for more massive systems the scatter in the simulated profiles somewhat exceeds that of the observed samples.

The FABLE simulations represent a major improvement over Illustris in the galaxy group and cluster regime. In particular, the baryonic content and global X-ray and SZ properties of the ICM are a good match to observations across a wide range of scales. Furthermore, the ICM is realistically distributed with residual deviations arising in the thermodynamic properties only toward the cluster centre. Our results are consistent with numerous other simulation studies and suggest that a subtle interplay between AGN feedback and a number of supplementary physical phenomena may be needed to explain the observed properties of galaxy clusters and groups in the core and the cluster outskirts.

3 THE REDSHIFT EVOLUTION OF X-RAY AND SZ SCALING RELATIONS

In this chapter I investigate the redshift evolution of the X-ray and Sunyaev-Zel'dovich (SZ) scaling relations for galaxy groups and clusters in the FABLE simulations. Using an expanded sample of 27 high-resolution zoom-in simulations, together with a uniformly-sampled cosmological volume to sample low-mass systems, I demonstrate very good agreement with the majority of observational constraints up to $z \sim 1$. The simulations predict significant deviations of all examined scaling relations from the simple self-similar expectations. While the slopes are approximately independent of redshift, the normalisations evolve positively with respect to self-similarity, even for commonly-used mass proxies such as the $Y_{\rm X}$ parameter. These deviations are due to a combination of factors, including more effective AGN feedback in lower mass haloes, larger binding energy of gas at a given halo mass at higher redshifts and larger non-thermal pressure support from kinetic motions at higher redshifts. These results have important implications for cluster cosmology from upcoming SZ surveys such as SPT-3G, ACTpol and CMB-S4, as relatively small changes in the observable–mass scaling relations (within theoretical uncertainties) have a large impact on the predicted number of high-redshift clusters and hence on our ability to constrain cosmology using cluster abundances. In addition, I find that the intrinsic scatter of the relations, which agrees well with most observational constraints, increases at lower redshifts and for lower mass systems. This calls for a more complex parametrization than adopted in current observational studies to be able to accurately account for selection biases.

The work presented in this chapter has been submitted for publication in MNRAS as:[1]

Henden N. A., Puchwein E., and Sijacki D., *The redshift evolution of X-ray and Sunyaev-Zel'dovich scaling relations in the FABLE simulations*.

[1]The analysis and discussion presented in this chapter is entirely my own work.

3.1 BACKGROUND

Ongoing and future clusters surveys with experiments such as SPT-3G, Advanced ACTpol, eROSITA and *Athena* hold great potential to provide precise constraints on the cosmological parameters (see Section 1.2.4). Yet this potential relies on our ability to relate the abundance of observed clusters, as a function of some observable, to theoretical predictions for the abundance of collapsed objects, as a function of their mass (Kravtsov & Borgani, 2012). Typically, cluster masses are inferred via the relationship between the total mass and an observable calibrated to a sample of clusters with more direct mass measurements, for example from gravitational lensing or an X-ray hydrostatic analysis. These mass–observable scaling relations are required to relate the theoretical mass function to the observed number counts and to understand the selection function of the survey, which describes how the observed cluster sample relates to the underlying population. Despite recent progress in the calibration of the mass–observable relations, the uncertainty in their slope and normalisation continues to dominate the error budget of current cosmological studies of clusters (e.g. Rozo et al. 2010; Sehgal et al. 2011; Mantz et al. 2015; Bocquet et al. 2015; Planck Collaboration XXIV 2016). Other aspects of the mass–observable relations are also not yet fully understood, for example, the origin of intrinsic scatter in the relations or their extension to low mass clusters or groups. Moreover, the mass-observable relations may vary with redshift beyond that expected in simple hierarchical models of cluster formation. As the increased size and depth of future surveys push cluster detections to increasingly high redshift, constraints on the redshift evolution of the cluster scaling relations will become increasingly important to exploit the full potential of these new samples.

The paucity of well-defined cluster samples at high redshift, and the lack of low mass clusters and galaxy groups in existing samples, strongly limits current constraints on the redshift evolution of the scaling relations. Existing observational studies (e.g. Vikhlinin et al. 2002; Maughan et al. 2006; Reichert et al. 2011; Hilton et al. 2012; Maughan et al. 2012; Sereno & Ettori 2015b) also find seemingly contradictory results. For example, some studies (e.g. Ettori et al. 2004; Reichert et al. 2011; Hilton et al. 2012) measure a negative evolution of the X-ray luminosity–temperature relation with respect to self-similarity, while others find zero or even positive evolution (e.g. Vikhlinin et al. 2002; Kotov & Vikhlinin 2005; Maughan et al. 2006; Pacaud et al. 2007). One of the dominant causes for this lack of consensus is the difficulty in accounting for selection bias in small, often heterogeneous samples of high-redshift clusters drawn from different surveys,

which can mimic evolution (e.g. Pacaud et al. 2007; Short et al. 2010).

Theoretical modelling of cluster formation can aid in understanding these issues by studying the evolution of cluster scaling relations for the same set of objects, or a well-defined subsample, over cosmic time. In the past decade, semi-analytic prescriptions have made significant progress in the modelling of realistic galaxy clusters (e.g. De Lucia & Blaizot 2007; Bower et al. 2008; Somerville et al. 2008; Guo et al. 2011; Contini et al. 2014), however these methods do not fully capture the effects of baryonic processes during cluster formation, which can have a significant impact on total halo masses (e.g. van Daalen et al. 2011; Cui et al. 2014b; Cusworth et al. 2014; Velliscig et al. 2014). An alternative approach is to use cosmological hydrodynamical simulations to follow the highly non-linear, dark matter-dominated growth of large-scale structure while, at the same time, self-consistently evolving the baryon component to make predictions for cluster observables.

A number of cosmological hydrodynamical simulations have been utilised to study the redshift evolution of the cluster scaling relations, with occasionally dissimilar results (e.g. Fabjan et al. 2011; Le Brun et al. 2016; Barnes et al. 2017a; Truong et al. 2018). Yet some variation in their predictions is to be expected given that the simulations use different sets of physical models with different parametrizations. For this reason it is important to explore a range of plausible models in order to constrain the dependence of the theoretical predictions on the physical modelling. In addition, observational constraints from future surveys have the potential to distinguish between models and thus to constrain the non-gravitational physics important to the formation and evolution of galaxy clusters.

In this chapter we explore the redshift evolution of X-ray and SZ scaling relations in the FABLE simulations. For this study we have performed an additional 21 high-resolution zoom-in simulations in addition to the initial six zoom-in simulations studied in Chapter 2. Here we extend the analysis of the $z = 0$ scaling relations in Chapter 2 out to $z \approx 2$ using our expanded sample.

This chapter is organised as follows. Section 3.2 describes our methods for calculating observable quantities and our choice of sample selection. In Section 3.3 we compare the X-ray scaling relations with observations at intermediate to high redshifts and investigate the evolution of the relations out to $z = 1.8$ with comparison to the recent simulation studies of Barnes et al. (2017a) and Truong et al. (2018). In Section 3.4 we explore the redshift evolution of the scaling between the SZ signal and total mass, including a comparison to observed clusters at $z \lesssim 1$. We also investigate how different predictions for the relation can affect the predicted cluster counts for future SZ surveys.

Throughout this chapter we assume a Planck cosmology (Planck Collaboration XIII, 2016) with cosmological parameters $\Omega_\Lambda = 0.6911$, $\Omega_\mathrm{M} = 0.3089$, $\Omega_\mathrm{b} = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$ and $H_0 = 67.74$ km s^{-1} Mpc^{-1}.

3.2 METHODS

3.2.1 *Simulations*

For this study we have greatly expanded our sample of zoom-in simulations from just 6 to a total of 27. As before, the resimulated haloes were selected from the dark matter-only (3 h^{-1} Gpc)^3 Millennium-XXL simulation (Angulo et al., 2012) to be approximately logarithmically spaced over the mass range $10^{13} M_\odot \lesssim M_{500} \lesssim 3 \times 10^{15}\, M_\odot$ at $z = 0$. These additional zoom-in simulations were performed to the same resolution with the same gravitational softening lengths and cosmology as described in Section 2.2.1.

In addition to the main halo of each zoom-in simulation we also consider "secondary" friends-of-friends (FoF; Davis et al. 1985) haloes within the high-resolution region. We include in our sample any FoF halo that is not contaminated by low resolution dark matter particles within $5\, r_{500}$ at the given redshift. This ensures that the halo properties are unaffected by the zoom-in technique. Indeed, we do not find any evidence that the X-ray or SZ properties of these secondary haloes depend systematically on their distance from the main halo or from the edge of the high-resolution region, which can be non-spherical.

3.2.2 *Calculating X-ray properties*

We estimate bolometric X-ray luminosities and spectroscopic temperatures for our simulated haloes via the method described in Chapter 2 with two minor alterations: the addition of Galactic HI absorption to the spectra and the inclusion of the metallicity information of the gas. We include the effects of Galactic HI absorption on the spectrum via a WABS model in XSPEC with a column density of 5×10^{20} cm^{-2}. This improves the realism of our spectra, although its effect on the derived X-ray luminosity and temperature is small (less than 2 per cent). Whereas in Chapter 2 we assumed a constant metallicity of 0.3 times the solar value for simplicity, here we utilise the metallicity of the gas tracked by the simulations. This is achieved by summing APEC emission models (Smith et al., 2001) for a series of temperature *and* metallicity bins. The conclusions of Chapter 2 are unchanged by using this updated method, however in the present study we are concerned with the exact slope of the X-ray scaling relations, which can be sensitive to

low temperature systems ($\lesssim 2$ keV), where metal line emission is a significant contributor to the total X-ray luminosity.

We mimic observations with either the *Chandra* or *Athena* X-ray observatories by convolving the mock spectrum with an appropriate response function as described below (Section 3.2.2.1). We follow the standard practice of fitting a single-temperature APEC model to the spectrum in the energy range 0.5–10 keV for *Chandra* and 0.2–12 keV for *Athena*. We fix the redshift to the input value but leave the temperature, metallicity and normalisation free to vary during the fit. The spectroscopic temperature is thus the temperature of the best-fitting model and the bolometric X-ray luminosity is calculated from the model in the energy range 0.01–100 keV.

For each halo we calculate an X-ray spectrum for the gas within a circular aperture of radius r_{500} centred on the minimum of the gravitational potential, integrated along the length of the simulation volume. We use a projected rather than a spherical aperture as this is more akin to observations. However, we caution that the resultant spectrum can be biased by hot gas along the line of sight. We find that the X-ray luminosity can be boosted by as much as ~ 15 per cent by gas that lies in projection, although the effect on the spectroscopic temperature is small ($\lesssim 2$ per cent). We shall comment on the effect of switching to a spherical aperture in situations where projection has an appreciable effect on the derived X-ray scaling relation. We note that other quantities, such as the total mass, gas mass and mass-weighted temperature, are measured directly from the simulation within a *spherical* aperture of radius r_{500}.

As in Chapter 2 we exclude cold gas with a temperature below 3×10^4 K and gas above the density threshold for star formation. We also exclude this gas from the mass-weighted temperature calculation. In this work we make one further temperature cut on the gas that excludes very high temperature bubbles created by our relatively simple model for radio-mode AGN feedback. Excessively hot AGN-heated bubbles bias the derived X-ray luminosity and spectroscopic temperature in a few per cent of systems and occur when a particularly strong feedback event has been very recently triggered. In reality, AGN-driven bubbles are thought to be supported by non-thermal pressure and should only contribute to the X-ray temperature once thermalisation has occurred. We find that such gas can be reliably excluded by applying a temperature threshold of 4.0 times the virial temperature, $k_B T_{200} = G M_{200} \mu m_p / 2 r_{200}$. This greatly reduces the presence of outliers, which can otherwise bias the scatter inferred from the X-ray scaling relations. We note that for the vast majority of systems no gas is excluded by this choice of threshold.

We also calculate the X-ray analogue of the integrated SZ effect known as $Y_{\rm X}$. First introduced by Kravtsov et al. (2006), $Y_{\rm X}$ is equal to the product of the core-excised spectroscopic temperature and total gas mass and is considered a low-scatter mass proxy that is especially robust to the cluster dynamical state (e.g. Arnaud et al. 2007; Maughan 2007; Nagai et al. 2007b) and to the baryonic physics included in simulations (e.g. Short et al. 2010; Fabjan et al. 2011; Planelles et al. 2014 but see also Le Brun et al. 2014). To obtain $Y_{\rm X}$ we calculate the core-excised spectroscopic temperature within a projected annulus of inner radius $0.15\,r_{500}$ and outer radius of r_{500}.

3.2.2.1 *Choice of response function*

In Section 3.3.1.2 we compare the X-ray properties of our simulations to observational data and for this we employ the response function and effective area energy curve of the *Chandra* ACIS-I detector, which is commonly used in cluster X-ray studies. We adopt a very large exposure time of 10^7 seconds so that we are not limited by photon noise even in low-mass galaxy groups at $z = 1$. The fits are performed in the energy range 0.5–10 keV.

For studying the redshift evolution of the X-ray scaling relations in Section 3.3.2 we find that current X-ray observatories such as *Chandra* possess insufficient effective area at low energies to reliably measure group and cluster temperatures out to high redshift ($z \approx 2$). In particular, our tests have shown that the spectroscopic temperature can be biased high in situations where a significant proportion of the X-ray emission is redshifted below the energy range used for the spectral fit. For example, using the *Chandra* response and fitting the spectra in the 0.5–10 keV range, we find that the spectroscopic temperature is biased high compared with the mass-weighted temperature at $z \gtrsim 1$. This bias increases with increasing redshift and is larger for lower temperature systems because a larger fraction of the X-ray emission is redshifted below the minimum energy of the fit. The size of this effect is large enough that it dominates the redshift evolution of the slope and normalisation of our spectroscopic temperature-based relations at $z \gtrsim 0.6$. Lower (higher) values for the minimum energy of the fit show decreased (increased) bias, however the effective area of *Chandra* ACIS-I becomes negligible at $\lesssim 0.5$ keV so that values smaller than 0.5 keV have little effect.

For this reason we generate spectra out to $z \approx 2$ using the X-ray Integral Field Unit (X-IFU) on board the future *Athena* X-ray observatory (Barret et al., 2018), which will possess an order of magnitude larger effective area than *Chandra* over a wider 0.2–12 keV bandpass. For this we use response matrices and effective area energy curves produced for the so-called cost-constrainted configuration of *Athena*

as described in Barret et al. (2018). For the mock *Athena* observations we adopt a very long exposure of 10^8 seconds, which ensures that the derived spectroscopic temperature is converged with respect to the total photon count for the lowest temperature objects in our sample at $z \approx 2$ (~ 1 keV). Such an exposure time is clearly impractical however it allows us to present predictions of the X-ray scaling relations over an extended halo mass range out to high redshift.

3.2.3 *Fitting of cluster scaling relations*

For all scaling relations we relate the property Y to the property X with a best-fitting power-law of the form

$$Y = E(z)^{\gamma}\ 10^{A}\ \left(\frac{X}{X_0}\right)^{\beta}, \tag{3.1}$$

where A and β describe the normalisation and slope of the relation, respectively, and $E(z)^{\gamma}$ corresponds to the expected self-similar evolution of the normalisation (see e.g. Borgani & Kravtsov 2011). X_0 is the pivot point, which we set to $M_{500} = 2 \times 10^{14}\ M_{\odot}$, $T_{500} = 3$ keV or $M_{\rm gas,500} = 2 \times 10^{13}\ M_{\odot}$ for the total mass, temperature or gas mass, respectively. These are close to the average values of our sample (defined in Section 3.2.4) across all redshifts ($z \leq 1.8$).

We perform the fitting in log-space using the orthogonal BCES method described in Akritas & Bershady (1996), which is commonly used in observational studies (e.g. Pratt et al. 2009; Zhang et al. 2011; Maughan et al. 2012; Giles et al. 2016). We note that in our case of no measurement errors this method reduces to orthogonal regression (e.g. Isobe et al. 1990). We have repeated our analyses using two other common choices for the fitting procedure, namely the BCES(Y|X) method (Akritas & Bershady, 1996) and the Bayesian approach described in Kelly (2007). We confirm that these two methods yield identical values for the best-fitting parameters. The orthogonal BCES method yields marginally higher values for the slope, although the difference is less than 5 per cent and comparable to the uncertainties. The offset is systematic across redshift bins and has a negligible effect on the redshift evolution of the slope and normalisation of the relations.

We also investigate the intrinsic scatter about the best-fitting relation, which we compute following Tremaine et al. (2002) as

$$\sigma = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}[\log_{10}(Y_i) - \log_{10}(F(X_i))]^2}, \tag{3.2}$$

where N is the number of data points, as described by their position (X_i, Y_i) in the space of observables X and Y, and F is the best-fitting power-law relation. We

have confirmed that the scatter calculated in this way is in good agreement with the best-fitting value found via the approach of Kelly (2007).

We estimate confidence intervals on the best-fitting parameters via bootstrap resampling of the data. Specifically, we generate 10^4 resamples with replacement and obtain the best-fitting parameters for each resample. The confidence interval for each parameter is defined as the empirical quantiles of the bootstrap distribution of the parameter following the "basic" bootstrap method described in Davison & Hinkley (1997). Quoted uncertainties on the best-fitting parameters correspond to the 68 per cent confidence interval.

3.2.4 *Sample selection*

A single power law adequately describes the scaling of massive clusters, however, at lower masses the slope of the relation can change due to the influence of non-gravitational processes such as feedback, which have a larger impact on lower mass haloes with shallower gravitational potential wells. In FABLE we find that certain scaling relations – particularly those involving X-ray luminosity or temperature – show signs of a steepening in the regime of low-mass galaxy groups with masses $M_{500} \lesssim 3 \times 10^{13}\ M_{\odot}$ and average temperatures $T_{500} \lesssim 1$ keV, consistent with previous simulation results (e.g. Davé et al. 2008; Puchwein et al. 2008; Gaspari et al. 2014; Planelles et al. 2014) as well as some observational studies (e.g. Helsdon & Ponman 2000; Mulchaey 2000; Sanderson et al. 2003; Eckmiller et al. 2011; Bharadwaj et al. 2015; Kettula et al. 2015).

To ensure that the best-fitting power law relation is not biased by low-mass groups we include only haloes above a mass threshold of $M_{500} > 3 \times 10^{13}\ E(z)^{-0.5}\ M_{\odot}$ in our sample. When plotting the X-ray and SZ scaling relations in Sections 3.3 and 3.4 we distinguish objects that are included in the fit from those that are not with filled and open symbols, respectively. The redshift evolution of the lower mass threshold is parametrised by the factor $E(z)^{-0.5}$, which was chosen to be similar to that of an SZ-selected sample based on the results of the 2500 deg^2 SPT-SZ survey (see e.g. fig. 6 in Bleem et al. 2015). Applying an SZ-like selection allows us to maximise the size of our sample at high redshift so that we are able to robustly derive the best-fitting scaling relations in single redshift bins.

This mass threshold corresponds to a sample of 39 haloes at $z = 0$, 68 at $z = 1$ and 44 at $z = 1.8$. Note that we do not extend our analyses beyond $z = 1.8$ as the number of cluster-scale haloes with $M_{500} > 10^{14} M_{\odot}$ falls to one. Our sample is not as large as some recent simulation studies (e.g. Barnes et al. 2017a; Le Brun et al. 2016) because our simulations are run at comparatively high resolution, which

limits the number of simulations we are able to realistically perform. On the other hand, our high-redshift samples are still comparable in size to recent observational studies of local scaling relations (e.g. Zou et al. 2016; Giles et al. 2017; Nagarajan et al. 2018; Ge et al. 2019).

Some simulation studies choose to limit their sample to $M_{500} \gtrsim 10^{14} M_{\odot}$ (e.g. Barnes et al. 2017a; Truong et al. 2018) in order to avoid a possible break in the power law fitting of certain cluster scaling relations. Indeed, there is some observational evidence that the scaling relation between X-ray luminosity and total mass or temperature experiences a break at $M_{500} \sim 10^{14} M_{\odot}$ (~ 3 keV; e.g. Hilton et al. 2012; Maughan et al. 2012; Lovisari et al. 2015), although others find a more gradual shift in slope (e.g. Eckmiller et al. 2011; Bharadwaj et al. 2015; Kettula et al. 2015) or no change at all (e.g. Anderson et al. 2015; Zou et al. 2016; Babyk et al. 2018). We have tested the changes to our best-fitting scaling relations when restricting our sample to more massive haloes with $M_{500} > 10^{14}\, E(z)^{-0.5}\, M_{\odot}$ and in general find a slightly shallower slope and smaller intrinsic scatter. We opt for a relatively low mass threshold ($M_{500} > 3 \times 10^{13}\, E(z)^{-0.5}\, M_{\odot}$) to obtain robust statistics but comment in the text on those aspects of the scaling relations that are affected by choosing a more massive sample.

3.3 X-RAY SCALING RELATIONS

In this section we explore various mappings between the total mass, gas mass, bolometric X-ray luminosity, X-ray spectroscopic temperature and $Y_{\rm X}$, including comparisons with observations at intermediate- and high-redshift (Section 3.3.1) and the redshift evolution of the relations (Section 3.3.2). We consider five X-ray scaling relations: gas mass, $Y_{\rm X}$ and X-ray luminosity as a function of total mass and total mass and X-ray luminosity as a function of temperature. In Section 3.3.1 only we also consider the relation between X-ray luminosity and gas mass. The ordering of the variables in each scaling relation is arbitrary since our choice of fitting procedure minimises the squared *orthogonal* distances of the data points to the best-fitting relation.

3.3.1 *Comparison to observations at intermediate and high redshift*

In Fig. 3.1 and 3.2 we compare the X-ray scaling relations at $z = 0.4$ and $z = 1.0$ with observed samples of clusters of similar median redshift (grey symbols). Solid lines show the best-fitting power law relation for the sample defined in Section 3.2.4, as indicated by filled diamonds. Observational data based on

weak lensing mass measurements are distinguished from X-ray hydrostatic mass estimates with open and filled symbols, respectively. In this section we mimic *Chandra* ACIS-I observations as described in Section 3.2.2. Below we describe the observational data sets used for the comparison and discuss the results in Section 3.3.1.2.

3.3.1.1 *Observational data*

We compare extensively to results from the XXL-100-GC sample, which consists of the 100 brightest clusters in the XXL survey over the redshift range $0.05 < z < 1.1$ (Pacaud et al., 2016). For the total mass–temperature relation ($z = 0.4$ only) we compare to XXL-100-GC clusters with direct weak lensing mass estimates from Lieu et al. (2016), restricting the sample to clusters at $0.259 \leq z \leq 0.52$ with a median redshift of 0.41. For the X-ray luminosity and gas mass-based relation we compare to data from Giles et al. (2016) and Eckert et al. (2016) for which masses are estimated from the weak lensing calibrated total mass–temperature relation derived in Lieu et al. (2016). We choose XXL-100-GC clusters at $0.3 < z < 0.5$ and $0.91 < z < 1.05$ with median redshifts 0.39 and 0.99, respectively. We note that the spectroscopic temperatures of the XXL-100-GC were measured within a circular aperture of fixed radius 300 kpc, however, Giles et al. (2016) find no systematic difference between this temperature and the temperature measured within r_{500}.

Additional weak lensing-based data come from Mahdavi et al. (2013) for a sample of clusters in the redshift range $0.15 < z < 0.55$ with combined *Chandra* and *XMM-Newton* X-ray data. We restrict our comparison to clusters at $0.28 < z < 0.55$ with a median redshift of 0.40. We note that the weak lensing mass estimates for this sample were revised upwards in Hoekstra et al. (2015) by approximately 20 per cent on average, however, updated values for the X-ray quantities (due to the associated increase in r_{500}) are not available. We therefore compare to the published data from Mahdavi et al. (2013) and describe in the text how an increase in the mass estimates may affect our comparison.

For the $Y_{\rm X}$–total mass relation at $z = 0.4$ we also compare to Mantz et al. (2016b) who measure gas masses and core-excised temperatures for massive cluster sample with *Chandra* and *ROSAT* X-ray data. We restrict their sample to clusters with weak lensing mass measurements at $0.35 < z < 0.45$ and a median redshift of 0.40.

Maughan et al. (2008), Reichert et al. (2011) and Hilton et al. (2012) study the X-ray scaling relations for large samples of clusters out to high redshift using X-ray hydrostatic mass estimates. Maughan et al. (2008) analyse 115 clusters at $0.1 < z < 1.3$ with archived *Chandra* data. For the $z = 0.4$ and $z = 1$ comparisons we limit their sample to $0.30 < z < 0.50$ (median 0.40) and $0.83 < z < 1.24$ (median

0.96), respectively. For the $Y_{\rm X}$–total mass comparison at $z = 0.4$ we use a subset of their clusters at $0.4 \lesssim z \lesssim 0.46$ with direct X-ray hydrostatic mass estimates given in Maughan (2007). Reichert et al. (2011) combine numerous published data sets to study the evolution of the X-ray scaling relations out to $z \sim 1.5$. For the $z = 0.4$ and $z = 1$ samples we use clusters at $0.30 < z < 0.50$ and $0.90 \leq z < 1.11$ with median redshifts of 0.41 and 0.99, respectively. Hilton et al. (2012) measure the evolution of the X-ray luminosity–temperature relation out to $z \sim 1.5$ using 211 clusters from the *XMM* Cluster Survey (Mehrtens et al., 2012). For the $z = 0.4$ and $z = 1$ comparisons we restrict their sample to $0.30 \leq z \leq 0.50$ and $0.91 \leq z \leq 1.13$ with median values of 0.41 and 1.00, respectively.

At $z = 1$ we supplement our comparison with data from Bartalucci et al. (2017) and Dietrich et al. (2019). Bartalucci et al. (2017) study five clusters at $0.93 < z < 1.13$ detected via the SZ effect with gas mass estimates derived from combined *XMM-Newton* and *Chandra* data. Halo masses are estimated using the total mass–$Y_{\rm X}$ relation of Arnaud et al. (2010) assuming self-similar evolution. We compare our $Y_{\rm X}$–total mass relation at $z = 1$ to the high-redshift cluster sample studied in Dietrich et al. (2019) with weak lensing mass estimates from *Hubble Space Telescope* data (Schrabback et al., 2018) and X-ray data from *Chandra*. We use seven of their clusters at $0.87 \leq z \leq 1.13$ and extract individual values of $Y_{\rm X}$ and M_{500} as presented in their fig. 11.

3.3.1.2 *Comparison to observations*

Figure 3.1 and 3.2 reveal many of the same trends as our comparison to low-redshift cluster data presented in Chapter 2. At all three epochs we have good agreement with the observed gas mass–total mass, $Y_{\rm X}$–total mass, X-ray luminosity–total mass and X-ray luminosity–gas mass relations based on X-ray hydrostatic masses (solid grey symbols) over a wide mass range. On the other hand, the FABLE clusters lie on the upper end of the observed scatter in the X-ray luminosity–temperature relation. This implies that the X-ray luminosities or spectroscopic temperatures of the simulated clusters may be over- or underestimated, respectively. Which of these interpretations dominates depends on whether observations based on X-ray hydrostatic or weak lensing masses are used for the comparison.

It is clear from Fig. 3.1 and 3.2 that several of the observed scaling relations change significantly when using weak lensing mass estimates (open symbols) as opposed to X-ray hydrostatic masses (solid symbols). From Fig. 3.1 we see that the $z = 0.4$ total mass–temperature and X-ray luminosity–total mass relations based on weak lensing are offset in normalisation compared with data based on

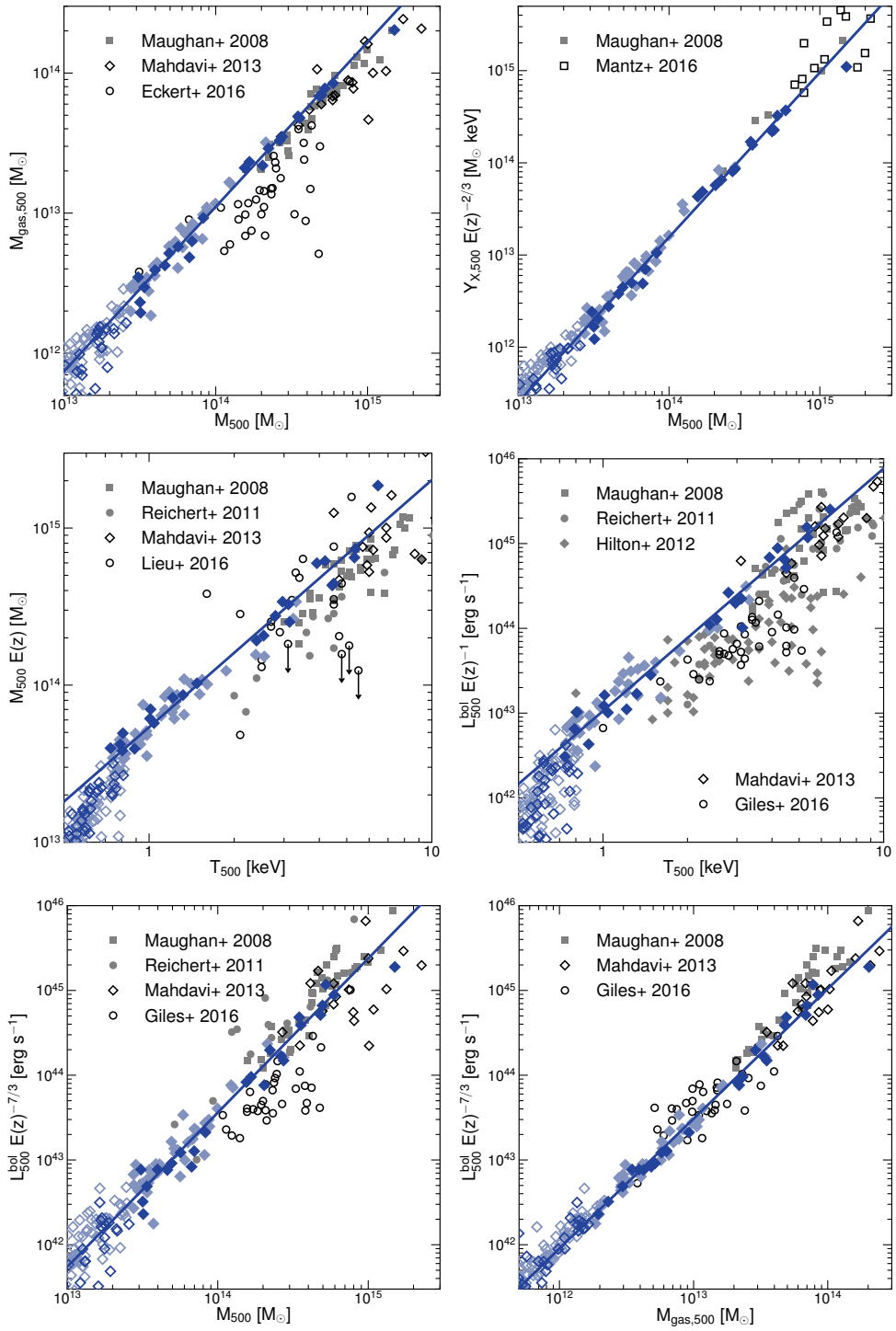


Figure 3.1: X-ray scaling relations at $z = 0.4$ (blue diamonds) compared with those derived from observations (grey symbols) at a similar median redshift as described in the text. Dark blue diamonds correspond to haloes in the full-volume simulation and the most massive halo in each zoom-in simulation, while light blue diamonds highlight lower mass companion clusters within the high resolution region of the zoom-in simulation. The solid line shows the best-fitting relation in each case and filled diamonds indicate the haloes that were included in the fit. Dark grey, open symbols represent observational data based on weak lensing while light grey, solid symbols correspond to X-ray only data.

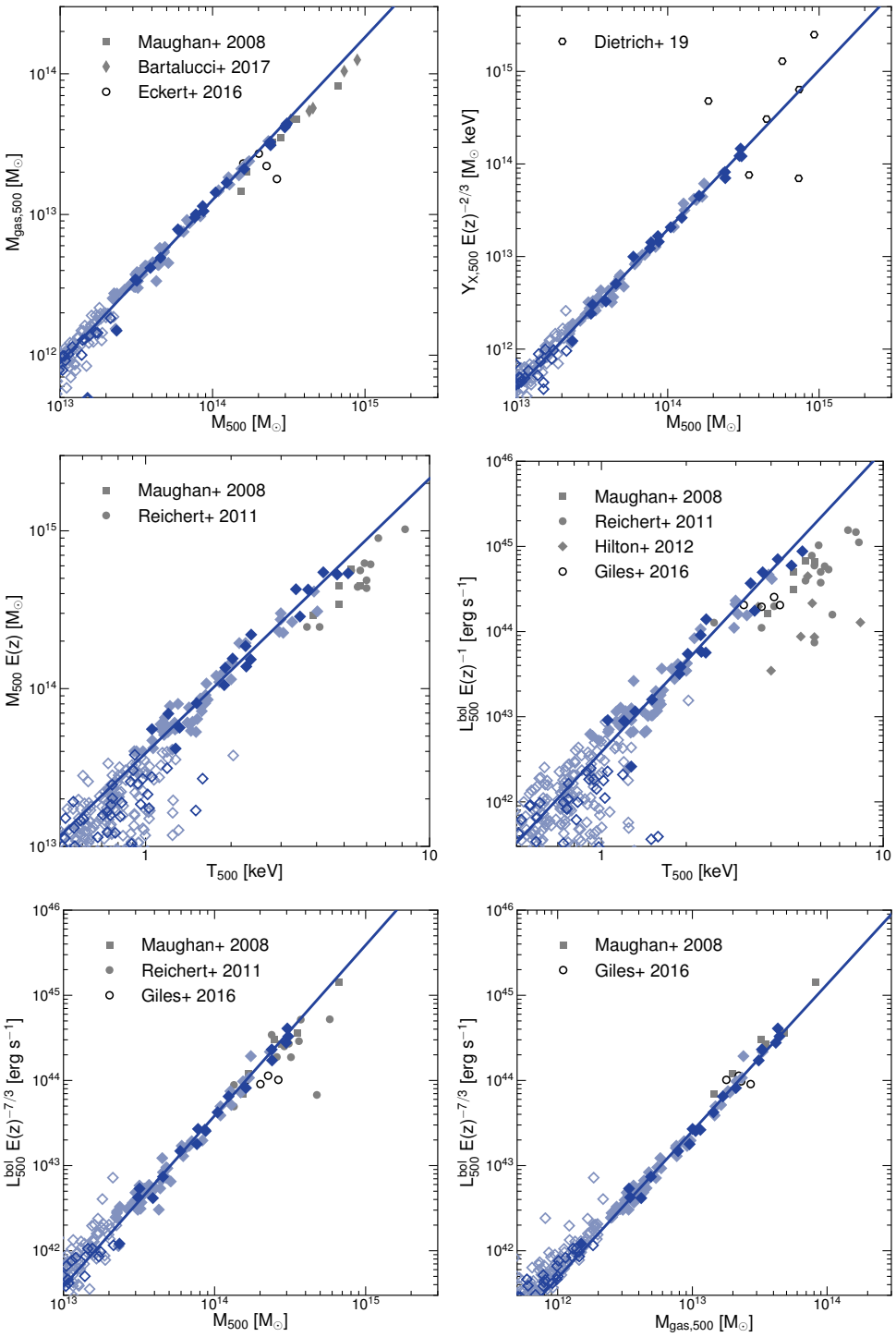


Figure 3.2: X-ray scaling relations at $z = 1.0$ compared with those derived from observations at a similar median redshift. Symbol styles are identical to those of Fig. 3.1.

X-ray masses, particularly if the Mahdavi et al. (2013) weak lensing masses are revised upwards as suggested by Hoekstra et al. (2015). It is unclear whether this difference continues to $z = 1$ due to the lack of weak lensing data at high redshift, although a similar offset in normalisation is found at $z \approx 0$ (see Section 2.5.5).

As for the $z \approx 0$ comparison in Chapter 2, whereas comparison of our relations to X-ray hydrostatic mass estimates would suggest that FABLE clusters tend to underestimate the temperature at fixed mass, comparison with (presumably less biased) weak lensing masses implies that the simulated clusters possess realistic global temperatures but are over-luminous at fixed mass. Given that the relation between X-ray luminosity and gas mass is a good match to the observations (bottom-right panels), the discrepancy in X-ray luminosity at fixed total mass must largely be driven by an overestimate in the gas mass. Indeed, the FABLE systems lie on the upper end of the scatter in the gas mass–total mass relation compared with the weak lensing-based studies of Mahdavi et al. (2013) and Eckert et al. (2016). As we discussed in Chapter 2, an increase in the efficiency of our AGN feedback model could reduce this discrepancy by ejecting larger gas masses from massive haloes, however, it is likely that a more sophisticated modelling of AGN feedback is required to simultaneously reproduce the cluster thermodynamic profiles.

Although the FABLE model tends to overpredict the gas masses and X-ray luminosities at fixed mass or temperature, the size of the offset does not change dramatically between $z = 0$ and $z = 1$. This gives us confidence that the FABLE model makes reliable predictions for the redshift evolution of the cluster scaling relations, which is the main topic of this chapter. Indeed, in the following section we investigate the redshift evolution of the X-ray scaling relations and show that the FABLE predictions are often bracketed by the results of two other recent simulation works that show different levels of agreement with observations.

3.3.2 *Evolution of the X-ray scaling relations*

In the following sections we assess the redshift evolution of the best-fitting parameters (slope, normalisation and intrinsic scatter) of each X-ray scaling relation. Here we mimic X-ray observations with the planned *Athena* X-IFU instrument as described in Section 3.2.2. We compare with recent results from the MACSIS (Barnes et al. 2017a; hereafter B17) and Truong et al. (2018) (hereafter T18) galaxy cluster simulations, which demonstrates how simulation predictions for the evolution of the X-ray scaling relations can vary given different choices for the physical modelling.

The MACSIS suite of zoom-in simulations consists of 390 galaxy clusters sim-

ulated with the baryonic physics model of the BAHAMAS simulation (McCarthy et al., 2017), which was calibrated to reproduce the present-day galaxy stellar mass function and the hot gas mass fractions of galaxy groups and clusters. The calibrated simulations reproduce a broad range of group and cluster properties at $z \approx 0$, including the X-ray and SZ scaling relations and the thermodynamic profiles of the ICM (McCarthy et al., 2017). B17 investigate the redshift evolution of a combined sample of clusters from the MACSIS and BAHAMAS simulations out to $z = 1.5$. Their sample consists of haloes above a redshift-independent mass limit of $M_{500} > 10^{14} M_{\odot}$, yielding 1294 clusters at $z = 0$ and 225 at $z = 1$. B17 calculate bolometric X-ray luminosities and spectroscopic temperatures from synthetic X-ray observations following Le Brun et al. (2014). We note that B17 use X-ray hydrostatic masses estimated from their mock X-ray data, which are biased low compared with the mass measured directly from the simulation (Henson et al., 2017). We caution that this may bias the slope of the observable–mass relations and their intrinsic scatter somewhat high compared with FABLE and T18. On the other hand, we do not expect a significant change to the redshift evolution of the slope, normalisation or intrinsic scatter of the MACSIS relations given that Henson et al. (2017) find no evidence for a redshift dependence of the X-ray mass bias.

T18 analyse a suite of 29 zoom-in simulations including AGN feedback that reproduce a wide range of cluster properties, most notably the observed dichotomy between cool-core and non-cool-core clusters (Rasia et al., 2015) and their thermodynamic profiles (Rasia et al., 2015; Planelles et al., 2017). Their zoom-in simulations are centred on clusters drawn from a $1\,h^{-3}$ Gpc^3 parent simulation (Bonafede et al., 2011) with a high-resolution Lagrangian region extending to at least five times the virial radius. Their cluster sample comprises all objects in the high-resolution regions above an evolving halo mass threshold of $M_{500} > 10^{14} E(z)^{-1} M_{\odot}$. T18 compute bolometric X-ray luminosities from pre-calculated cooling-function tables assuming the APEC model and approximate the spectroscopic temperature using the "spectroscopic-like" temperature (Mazzotta et al., 2004). Both T18 and B17 calculate their X-ray luminosities and temperatures within a spherical, core-excised aperture ($0.15 < r/r_{500} < 1$). We have repeated our analyses for the same aperture and comment on the change to the best-fitting parameters in the appropriate text.

We calculate the normalisation of the relations at the pivot points used in our own fitting procedure, which are $M_{500} = 2 \times 10^{14} M_{\odot}$ and $T_{500} = 3$ keV. These are very close to the pivot points used in T18 ($M_{500} \approx 1.5 \times 10^{14} M_{\odot}$ and $T_{500} = 3$ keV) although somewhat lower than MACSIS ($M_{500} = 4 \times 10^{14} M_{\odot}$ and $T_{500} = 6$ keV; B17). We propagate the uncertainty in the normalisation assuming that the

uncertainties on the best-fitting slope are independent of the normalisation and normally distributed. These assumptions are not necessarily valid for MACSIS and T18, however we find no systematic bias in the uncertainties when applying the same procedure to our own uncertainties for a range of different pivot points. To enable a comparison of the redshift evolution of the normalisation we plot the normalisation at each redshift relative to the $z = 0$ value. Because the best-fitting scaling relations include a term for self-similar evolution of the normalisation (Section 3.2.3), any deviation of the curves from horizontal indicates departure from self-similarity. Positive (negative) evolution refers to a normalisation that increases (decreases) with increasing redshift relative to the self-similar expectation.

B17 and T18 calculate the intrinsic scatter as in equation 3.2 except that B17 take N rather than $N - 2$ in the denominator. Their sample is large enough however that the difference is negligible. When quoting the intrinsic scatter measured in other studies we convert to units of dex to be consistent with the definition in equation 3.2.

We also make frequent reference to the simulation study of Le Brun et al. (2016) although we do not plot their best-fitting parameters. Le Brun et al. (2016) analyse the cosmo-OWLS suite of simulations (Le Brun et al., 2014), which employ four different galaxy formation models. Unless otherwise stated we refer to their fiducial AGN 8.0 simulation, which Le Brun et al. (2014) have shown produces the best match to observations. The AGN 8.0 model is similar to that of BAHAMAS and MACSIS except for slight adjustments to the parameters of the stellar and AGN feedback models (see table 1 in McCarthy et al. 2017). Le Brun et al. (2016) construct a sample of all haloes with $M_{500} > 10^{13} M_{\odot}$ at each redshift up to $z = 1.5$. They fit their scaling relations using both a single power law and a broken power law with low- and high-mass slopes below and above the pivot point ($M_{500} = 10^{14} M_{\odot}$). We refer to their single power law fit unless stated otherwise.

In the following sections we discuss the X-ray scaling relations in turn. We focus much of our attention on interpreting the gas mass–total mass and total mass–temperature relations (Sections 3.3.2.1 and 3.3.2.2) as the other relations are closely related to these.

3.3.2.1 *Gas mass–total mass scaling relation*

Figure 3.3 shows, from left to right, the redshift evolution of the best-fitting slope, normalisation and intrinsic scatter of the gas mass–total mass relation, respectively.

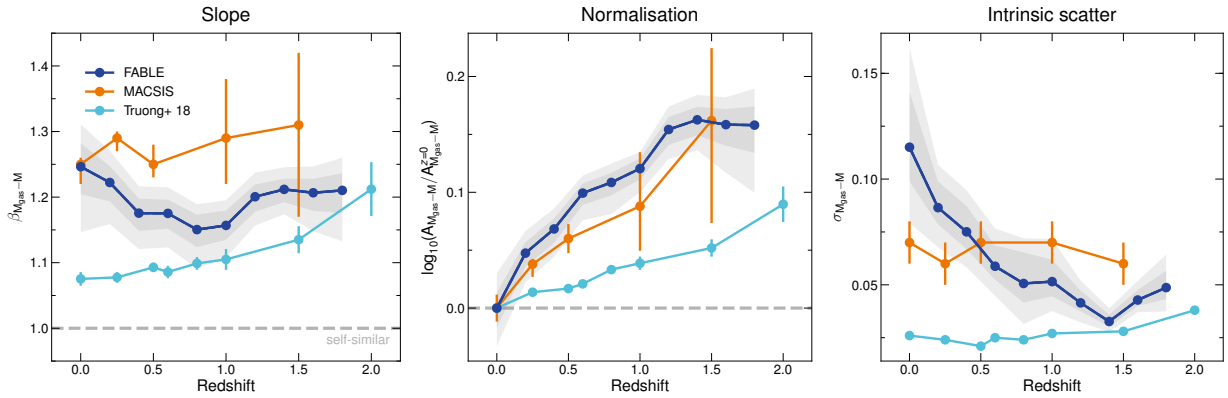


Figure 3.3: The redshift evolution of the gas mass–total mass relation for the FABLE (dark blue), MACSIS (orange) and T18 (light blue) simulations. The panels show, from left to right, the slope, normalisation and intrinsic scatter of the best-fitting power law relation as a function of redshift. Shaded regions about the FABLE relation represent the 68 and 95 per cent confidence intervals on the best-fitting parameters estimated from bootstrap resampling. The uncertainties on the intrinsic scatter are unknown for the T18 relation. Grey dashed lines indicate the self-similar expectation.

Gas mass–total mass slope: The self-similar expectation for the slope of the gas mass–total mass relation is unity (i.e. a constant gas mass fraction). This is indicated by a horizontal dashed line in the left-hand panel of Fig. 3.3. FABLE, MACSIS and T18 predict a gas mass–total mass slope significantly greater than unity out to $z \sim 2$. This is consistent with the results of Chiu et al. (2016a, 2018) who find a steeper than self-similar mass trend in the gas content of galaxy groups and clusters out to $z \approx 1.3$ with no significant redshift dependence. Previous numerical studies have shown that simulations incorporating radiative cooling and star formation yield a steeper than self-similar gas mass–total mass relation due to the conversion of gas into stars, which occurs more efficiently in lower mass systems (e.g. Stanek et al. 2010; Battaglia et al. 2013; Planelles et al. 2014; Le Brun et al. 2016). In addition, simulations that include some form of AGN feedback produce an even steeper slope as feedback is able to expel gas more efficiently from lower mass haloes due to their shallower potential wells, leading to a tilt in the relation (e.g. Puchwein et al. 2008; Fabjan et al. 2011; McCarthy et al. 2011; Gaspari et al. 2014).

At $z = 0$, FABLE and MACSIS predict a slope of $1.25^{+0.04}_{-0.04}$ and $1.25^{+0.01}_{-0.03}$, respectively. These values are in good agreement with a number of observational constraints, including Arnaud et al. (2007) (1.25 ± 0.06), Gonzalez et al. (2013) (1.26 ± 0.03), Eckert et al. (2016) ($1.21^{+0.11}_{-0.10}$) and Chiu et al. (2018) (1.32 ± 0.07). T18 derive a significantly shallower slope of 1.08 ± 0.01 at $z = 0$, although still steeper

than self-similar. Some observational studies measure a similarly shallow slope, for example Lin et al. (2012) (1.13 ± 0.03), Mahdavi et al. (2013) (1.04 ± 0.10) and Mantz et al. (2016a) (1.04 ± 0.05), the latter two results being consistent with self-similarity. Differences between studies can be attributed to a number of factors. One of the most influential of these is the sample selection, such as the distribution of mass, redshift and dynamical state of the clusters. For instance, the sample of Mantz et al. (2016a) is comprised of massive relaxed clusters that, due to their deep potential wells, have lost a comparatively small fraction of their gas through AGN feedback. We also find a mass dependence in the gas mass–total mass slope. For example, if we restrict our $z = 0$ sample to haloes with $M_{500} > 10^{14} M_{\odot}$ as in the MACSIS and T18 samples then the best-fitting slope drops to $1.16^{+0.05}_{-0.05}$, which lies in between their results.

The simulation results additionally depend on the model implementation of AGN feedback. The FABLE, MACSIS and T18 models all inject AGN feedback energy thermally, however they differ as to when and how much energy is input. In both FABLE and MACSIS the frequency of energy injection is controlled by a duty cycle, although in MACSIS the duty cycle depends on the ability of the AGN to heat the surrounding gas (Booth & Schaye, 2009; McCarthy et al., 2017), while in FABLE it is controlled either by a fixed accumulation time or by the mass growth of the black hole, depending on whether the AGN is in the quasar- or radio-mode (see Section 2.2.4 and Sijacki et al. 2007). Conversely, the T18 model for AGN feedback inputs feedback energy continuously (Steinborn et al., 2015). Continuous AGN feedback may have a gentler impact on the ICM compared with the duty cycle models of FABLE and MACSIS, which store feedback energy before injecting it into the surrounding gas in a single energetic event. For example, a number of studies have found that more energetic but less frequent AGN feedback events are more effective at reducing the gas content of galaxy groups and clusters (e.g. Le Brun et al. 2014, Schaye et al. 2015 and Appendix A). Since AGN feedback is more effective in lower mass haloes, this may explain why the FABLE and MACSIS models predict a slightly steeper gas mass–total mass relation compared with T18.

Although the AGN feedback models used in FABLE and MACSIS are similar in the sense that the feedback energy is input thermally and on a duty cycle, the slope of the gas mass–total relation in FABLE is somewhat shallower than MACSIS at most redshifts. This suggests that the removal of gas via AGN feedback is less efficient in FABLE than in MACSIS. Indeed, in Section 3.3.1.2 we showed that FABLE haloes are seemingly too gas-rich at fixed halo mass, implying the need for more efficient feedback. Some of the difference in slope can also be attributed to X-ray mass bias in the MACSIS results, which Henson et al. (2017) show increases mildly

with mass, thereby steepening the gas mass–total mass relation.

At $z \lesssim 1$ the gas mass–total mass slope decreases mildly with increasing redshift. This is consistent with Le Brun et al. (2016) who predict a decrease in slope with increasing redshift out to $z = 1.5$ for their simulations with AGN feedback. This evolution can partly be attributed to the reduced efficiency of gas expulsion by AGN feedback with increasing redshift, which we discuss in more detail in the next section. In addition, the change in slope at $z \lesssim 1$ corresponds to an increase in the proportion of central black holes operating in the radio-mode of AGN feedback (from a constant ~ 90 per cent at $z \gtrsim 1$ to almost 100 per cent at $z = 0$ in our sample), which occurs when the black hole accretion rate drops below one per cent of the maximal Eddington rate. In Appendix A we show that the radio-mode of feedback is more efficient than the quasar-mode at reducing gas mass fractions. Therefore, the increasing prevalence of the radio-mode with decreasing redshift may partly explain the increase in the gas mass–total mass slope. Similarly, we find a rapid change in the normalisation of the gas mass–total mass relation at $z \lesssim 1$ (middle panel of Fig. 3.3) consistent with increased gas expulsion via radio-mode AGN feedback.

At $z \gtrsim 1$ the slope is approximately independent of redshift. In fact, within the uncertainties the evolution in the slope is consistent with zero over a much wider redshift range ($0.4 \lesssim z \leq 1.8$), which is consistent with MACSIS and with the simulations with AGN feedback in Fabjan et al. (2011) and Battaglia et al. (2013). T18 find a positive evolution in the slope at $z \gtrsim 1$, however they attribute this to the decreasing mass of their sample with increasing redshift coupled with gas mass fractions that decline towards lower mass haloes due to the aforementioned effects of AGN feedback. This effect is accentuated for their sample compared with FABLE and MACSIS as their minimum mass threshold falls more rapidly with redshift. These results suggest that, whilst differences in the theoretical modelling lead to slightly different predictions for the slope of the gas mass–total mass relation, the change in slope with redshift is typically small in comparison.

Gas mass–total mass normalisation: In the middle panel of Fig. 3.3 we plot the normalisation of the gas mass–total mass relation as a function of redshift. All three simulations yield a positive evolution in normalisation relative to the self-similar model, which predicts no evolution (i.e. constant gas fraction with redshift). This implies either that AGN feedback is less effective at expelling gas at high redshift or that the efficiency with which gas is cooled and converted into stars decreases with increasing redshift. In FABLE it does not seem that the latter explanation holds since the total stellar mass within r_{500} at fixed total mass shows

little evolution at $z < 2$ (not shown), consistent with the results of an SZ-selected sample at $0.2 < z < 1.25$ (Chiu et al., 2018). The evolution of the normalisation must therefore be driven by AGN feedback. At high redshift, clusters of a given mass are denser and have deeper gravitational potential wells. This increases the binding energy of the halo so that AGN feedback must supply more energy in order to eject gas beyond r_{500}. Furthermore, central AGN at high redshift have had less time in which to affect the ICM of their host cluster. These effects combine to reduce the normalisation of the gas mass–total mass relation with decreasing redshift.

T18 predict a slightly weaker evolution than FABLE and MACSIS, which suggests that hot gas removal by cooling, star formation and AGN feedback is slightly less efficient in their simulations, at least at $z \lesssim 2$. We note that the difference between FABLE and T18 is reduced by using a higher pivot point since the slope of the relations behaves differently with redshift, however the offset remains significant for any reasonable choice of pivot point. Contrary to these results, simulations works such as Planelles et al. (2013) and Battaglia et al. (2013) find a constant gas and baryon mass fraction with redshift up to $z = 1$, which suggests that feedback is not removing gas at all in these simulations at $z < 1$.

Gas mass–total mass intrinsic scatter: The three simulations predict quite different levels of intrinsic scatter at low redshift. For instance, at $z = 0$ the measured scatter is $0.12^{+0.03}_{-0.02}$ in FABLE, 0.07 ± 0.01 in MACSIS and 0.026 in T18. Observations typically measure an intrinsic scatter close to $\sigma_{\mathrm{M_{gas}}} \approx 0.05$, for example, Arnaud et al. (2007) (0.044), Mahdavi et al. (2013) (0.07 ± 0.03), Mantz et al. (2016a) (0.04 ± 0.01) and Chiu et al. (2018) (0.05 ± 0.01). These constraints are lower than the FABLE prediction at low redshift but slightly higher than T18.

Some of the variation can be attributed to sample differences, in particular the mass range. Indeed, less massive objects tend to exhibit larger scatter in their X-ray properties, as has been shown for the cosmo-OWLS, BAHAMAS and MACSIS simulations (Le Brun et al., 2016; Farahi et al., 2018). Similarly, Eckmiller et al. (2011) measure significantly increased intrinsic scatter for a sample of 26 local galaxy groups compared with a more massive sample of 64 clusters from the HIFLUGCS survey (Hudson et al., 2010). As our sample is significantly less massive than that of MACSIS, T18 and the aforementioned observational studies, our scatter measurement is likely biased high by low-mass objects. Indeed, we find that increasing the mass of our sample significantly lowers the intrinsic scatter at all redshifts. For instance, at $z = 0$ the intrinsic scatter drops from $0.12^{+0.03}_{-0.02}$ to $0.07^{+0.03}_{-0.01}$ when restricting our sample to the same mass range as MACSIS and T18

at $z = 0$ ($M_{500} \geq 10^{14} M_{\odot}$). This brings the scatter into agreement with MACSIS and most observational constraints.

The increase in intrinsic scatter towards lower masses is likely associated with the increasing influence of non-gravitational processes such as stellar and AGN feedback in less massive haloes. Indeed, numerical studies such as Stanek et al. (2010) and Le Brun et al. (2016) have shown that preheating or AGN feedback not only increases the intrinsic scatter of the gas mass at fixed halo mass but also strengthens the trend of increasing scatter with decreasing halo mass (see e.g. fig. 9 in Stanek et al. 2010). To confirm this result in our simulations, we have re-simulated our $(40\,h^{-1}\,\mathrm{Mpc})^3$ volume without AGN feedback. We find that the FABLE model has increased intrinsic scatter compared with the non-AGN run for all of the scaling relations presented here and, although the volume is limited to haloes with $M_{500} \lesssim 10^{14} M_{\odot}$, there is a distinct trend of increasing scatter towards lower halo masses in all scaling relations.

T18 predict an intrinsic scatter two to three times smaller than FABLE and MACSIS at $z = 0$ for the same mass threshold, $M_{500} \geq 10^{14} M_{\odot}$. T18 expect their instrinsic scatter to be biased low due to their small sample size, which limits the number of outliers, yet their sample is three times larger than ours for the same mass range. Part of this offset may be explained by differences in the AGN feedback modelling, in particular the magnitude and frequency of the thermal energy injections. As mentioned above, the AGN feedback models of FABLE and MACSIS operate on a duty cycle while the T18 model inputs energy continuously. With a duty cycle, a large amount of feedback energy can be input to the ICM in one event. This can have a sudden and significant impact on ICM properties such as the gas mass and temperature, which we confirm in our model. The current ICM properties can therefore vary depending on the time that has passed since the last feedback event. As a result we might expect AGN feedback models with a duty cycle to produce stronger or more numerous outliers than a continuous feedback model, yielding a larger intrinsic scatter.

We find that the intrinsic scatter decreases with increasing redshift in FABLE, in contrast to T18 and MACSIS which predict little to no evolution. This may be driven by the increasing prevalence of radio-mode AGN feedback with decreasing redshift in our simulations as mentioned above. Since the radio-mode is more effective than the quasar-mode at removing gas from massive haloes, a higher fraction of feedback occurring in the radio mode at low redshift may introduce a larger scatter in the gas mass at fixed total mass. Furthermore, the intrinsic scatter may decrease with increasing redshift as AGN feedback in either mode becomes less effective at depleting halo gas mass fractions due to their increased

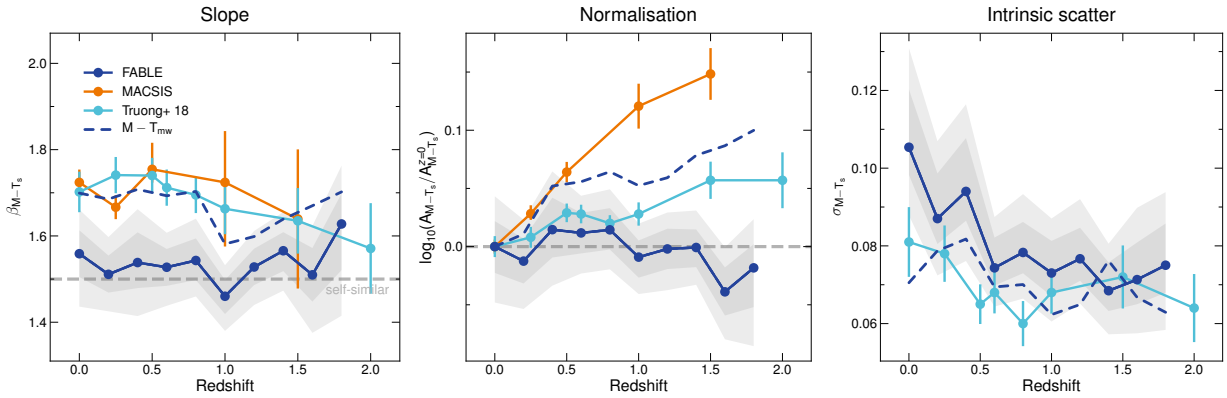


Figure 3.4: As Fig. 3.3 showing the redshift evolution of the total mass–temperature relation. The left, middle and right-hand panels show the slope, normalisation and intrinsic scatter, respectively. Blue solid and dashed lines correspond to the FABLE relation calculated using the spectroscopic temperature and mass-weighted temperature, respectively. The intrinsic scatter in the MACSIS relation is not shown as B17 compute their scatter at fixed mass rather than fixed temperature.

binding energy at fixed mass. We point out that significant redshift evolution in the intrinsic scatter could have important implications for cluster cosmology, which requires knowledge of the intrinsic scatter in order to properly account for selection biases (e.g. Maughan et al. 2012).

3.3.2.2 *Total mass–temperature scaling relation*

In Fig. 3.4 we plot the redshift evolution of the best-fitting parameters of the total mass–spectroscopic temperature relation. We also show with a dashed line the parameters of the total mass–temperature relation based on the mass-weighted temperature. We do not plot the uncertainties on these parameters but they are comparable to those of the spectroscopic temperature relation. The spectroscopic temperature is closer to that measured in X-ray observations, however it can be biased with respect to the mass-weighted temperature, which is a direct measure of the total thermal energy of the ICM.

Total mass–temperature slope: The slope of the total mass–temperature relation based on the mass-weighted temperature (dashed line) is steeper than self-similar at all redshifts (i.e., $\beta_{\rm M-T} > 1.5$), in good agreement with MACSIS and T18. The same departure from self-similarity was also found in the numerical studies of Stanek et al. (2010), Fabjan et al. (2011), Pike et al. (2014) and Le Brun et al. (2016) up to $z \sim 1$.

There are several ways in which the physical processes included in these simulations can affect the average temperature of the ICM, which in the self-

similar scenario is determined solely by the depth of the gravitational potential well. For example, radiative cooling can cool the dense gas to form stars, thereby removing low entropy gas and raising the average temperature of the hot phase. This occurs with greater efficiency in lower mass systems, resulting in a tilt in the total mass–temperature relation. Furthermore, AGN feedback can raise the average entropy of the ICM, particularly in the cluster core, leading to a higher average temperature. Indeed, comparing with our simulation that repeats the $(40\,h^{-1}\,\mathrm{Mpc})^3$ periodic volume but without AGN we find on average a small increase (~ 0.05 dex) in the spectroscopic and mass-weighted temperature when including AGN feedback. We lack enough massive haloes in this volume to constrain the slope of the total mass–temperature relation, however previous numerical studies have shown that AGN feedback plays a role in driving the slope away from the self-similar expectation (e.g. Fabjan et al. 2011; Pike et al. 2014; Le Brun et al. 2016).

The majority of observational studies also measure a steeper than self-similar slope, for example, Arnaud et al. (2007) (1.71 ± 0.09), Reichert et al. (2011) (1.76 ± 0.08) and Lieu et al. (2016) (1.67 ± 0.12). Although, others studies yield results consistent with self-similarity, for example Kettula et al. (2015) ($1.52^{+0.17}_{-0.16}$) and Mantz et al. (2016a) (1.52 ± 0.11). Observed cluster samples are typically limited to objects at or below $z \sim 1$ where the slope predicted by MACSIS, T18 and the mass-weighted FABLE relation agree on a roughly redshift-independent value of $\beta_{\mathrm{M-T}} \sim 1.7$, in good agreement with the majority of the observations.

The slope of the total mass–spectroscopic temperature relation (solid line) is significantly shallower than the mass-weighted temperature relation (dashed line) at all redshifts and is consistent with self-similarity. We have investigated this discrepancy and find that the spectroscopic temperature is biased slightly low compared with the mass-weighted temperature in galaxy groups ($T \lesssim 2$ keV) and biased somewhat high in massive clusters ($T \gtrsim 5$ keV).

At $z = 0$ the spectroscopic temperature is approximately 0.1 dex lower than the mass-weighted temperature at $\lesssim 2$ keV. At these temperatures we find that the mock X-ray spectra are poorly fit by a single-temperature model, which is biased towards the lower temperature component(s) of the spectrum and tends to underestimate the X-ray continuum emission at high energies. Mazzotta et al. (2004) find the same qualitative result in two-temperature thermal spectra for which the lower-temperature component is $\lesssim 2$ keV. Indeed, we find that a two-temperature model is a significantly better fit to our mock spectra, as found by de Plaa et al. (2017) for a large sample of clusters, groups and elliptical galaxies observed with *XMM-Newton*. We have used a single-temperature fit for consistency with most

observational constraints on the total mass–temperature relation but caution that this can underestimate the mass-weighted temperature in the galaxy group regime. This causes a significant tilt in our best-fitting total mass–spectroscopic temperature relation due to the high proportion of galaxy groups in our sample. We find a similar level of bias for simulated *Chandra* observations with realistic exposure times (10 ks), which suggests that current observational constraints may also be affected. Indeed, studies of the total mass–temperature relation in the galaxy group regime measure a slope slightly higher than, but statistically consistent with, our $z = 0$ result ($1.56^{+0.06}_{-0.06}$), for example, Sun et al. (2009) (1.65 ± 0.04), Eckmiller et al. (2011) (1.68 ± 0.20), Kettula et al. (2013) ($1.48^{+0.13}_{-0.09}$) and Lovisari et al. (2015) (1.65 ± 0.07), all of which derive the spectroscopic temperature from a single-temperature fit. It is possible that the bias is caused by an excess of cool, X-ray emitting gas in our simulated galaxy groups, however this is difficult to constrain from observations. Our tests have shown that, if such gas is present in our simulations, it is neither gravitationally bound in substructures nor does it belong to the separated cooling phase of gas identified in the cluster simulations of Rasia et al. (2012) (see their appendix A).

We can avoid the bias at low temperatures by limiting our sample to higher masses, however, the mass–temperature slope remains at a similarly low value due to an opposite spectroscopic temperature bias at the high mass end. In Section 2.6 we showed that the temperature and entropy profiles of FABLE clusters show signs of over-heating in the central regions due to our relatively simple model for radio-mode AGN feedback. Because the density, and thus the X-ray emissivity, of the ICM increases towards the cluster centre, this causes the spectroscopic temperature to be biased high relative to the mass-weighted temperature. Indeed, we find that excising the cluster core ($r < 0.15\, r_{500}$) from the temperature computation largely removes the spectroscopic temperature bias at the high mass end. If we also avoid the bias at the low mass end by restricting our sample to higher temperatures ($\gtrsim 2$ keV), we find that the slope of the total mass–spectroscopic temperature relation is in good agreement with the mass-weighted one. We note that a more sophisticated model for AGN feedback may address the spectroscopic temperature bias at the high mass end by bringing the thermodynamic profiles of our simulated clusters into better agreement with observations in the central regions.

FABLE and MACSIS predict a roughly redshift-independent total mass–temperature slope within the uncertainties. In contrast, T18 predict a mild decrease in the slope with redshift at $z \gtrsim 1$. They attribute this evolution to relatively cool gas at high redshift that has yet to thermalise in low-mass objects, whereas the most massive systems have been heated by strong shocks driven by minor and

major mergers. This could indicate more intense AGN activity at high redshift in FABLE and MACSIS compared with T18, which would raise the temperature of the gas preferentially in the lowest mass systems and cause a steepening of the total mass–temperature relation with increasing redshift. Indeed, T18 show this to be the case in their simulations with and without AGN feedback.

Total mass–temperature normalisation: The middle panel of Fig. 3.4 shows the normalisation of the total mass–temperature relation as a function of redshift. In FABLE we find a mild positive evolution in the normalisation when using the mass-weighted temperature (dashed line). This implies that objects of a given mass at high redshift are cooler than expected from the self-similar model. In contrast, the evolution of the spectroscopic temperature-based relation is consistent with the self-similar prediction. The difference between this and the mass-weighted temperature relation is due to redshifting of low-energy X-ray emission below the X-ray bandpass (0.2–10 keV), which causes the spectroscopic temperature to be biased high. As discussed in Section 3.2.2.1, the size of the bias is dependent on the minimum energy of the bandpass. For example, using the *Chandra* response function and fitting the spectra in the 0.5–10 keV range actually causes the normalisation to evolve negatively. We have confirmed this effect by fitting our mock X-ray spectra in the rest frame of the source, in which case the evolution of the normalisation is almost identical to the mass-weighted case.

All three simulations predict a positive evolution in the normalisation, although it is somewhat stronger in MACSIS (e.g. an increase of ~ 35 per cent from $z = 0$ to $z = 1.5$ compared with an increase of ~ 15 per cent for T18 and the mass-weighted temperature relation of FABLE). The difference may be driven by the slight redshift dependence of the spectroscopic temperature bias in MACSIS clusters, as demonstrated in fig. 7 of B17. They attribute this temperature bias to relatively cool X-ray emitting gas in the outskirts of massive clusters, although it is unclear whether this gas has a physical origin or is an unphysical artefact of the SPH hydrodynamics scheme (Henson et al., 2017). Interestingly, Le Brun et al. (2016) also find a positive evolution in the normalisation even in a simulation without radiative cooling, star formation or feedback. This implies that the redshift evolution is driven by the merger history of clusters rather than non-gravitational physics. Indeed, Le Brun et al. (2016) find that the ratio of the total kinetic energy of the ICM to the total thermal energy strongly increases with increasing redshift for a given halo mass due to the increasing importance of mergers and associated lack of thermalisation. We have confirmed the same redshift trend of the kinetic-to-thermal energy ratio in FABLE (not shown), as do T18. This implies that clusters

of a given mass possess greater non-thermal pressure support from bulk motions and turbulence at higher redshift, resulting in a lower temperature required for virial equilibrium.

The fact that simulation studies find a positive evolution in the total mass–temperature normalisation regardless of the precise physical modelling implies that this prediction is fairly robust. Yet this appears to be in mild tension with the results of Reichert et al. (2011) who use observational data from various literature sources to show that the evolution of the total mass–temperature relation is consistent with the self-similar prediction out to $z \sim 1.4$. This is similar to our spectroscopic temperature relation, which suggests that part of the discrepancy may be the result of a redshift-dependent spectroscopic temperature bias similar to that found in our mock X-ray analysis. Further observational constraints on the redshift evolution of the total mass–temperature normalisation are required with which to compare the simulation predictions however, particularly as the Reichert et al. (2011) result may be adversely affected by sample selection biases, which are of particular concern in imhomogeneous datasets drawn from multiple sources.

The *Athena* X-ray observatory will provide an excellent opportunity to constrain the biases in previous analyses and to test the simulation predictions (Nandra et al., 2013; Barcons et al., 2017). Indeed, our mock *Athena* X-IFU observations suggest that *Athena*, with careful consideration of a possibly redshift-dependent spectroscopic temperature bias, should observe a significant positive evolution in the total mass–temperature relation out to $z \sim 2$. The size of this evolution (or potentially its sign) will place constraints on the non-gravitational physics important in galaxy cluster formation and provide an opportunity to distinguish between different physical models.

Total mass–temperature intrinsic scatter: The intrinsic scatter about the total mass–temperature relation at $z = 0$ is $0.10^{+0.02}_{-0.01}$ and $0.07^{+0.01}_{-0.01}$ dex for the spectroscopic and mass-weighted temperature relations, respectively. These values are consistent with the observational studies of Eckmiller et al. (2011) (0.117), Kettula et al. (2013) (0.12 ± 0.03) and Kettula et al. (2015) (0.07 ± 0.04) but smaller than Lieu et al. (2016) (0.18 ± 0.03) and larger than the relaxed cluster samples of Arnaud et al. (2007) (0.064) and Mantz et al. (2016a) (0.058 ± 0.008). The T18 results are consistent with ours within the uncertainties. The intrinsic scatter at fixed mass is also in good agreement with MACSIS (not shown). Interestingly, we find that excising the core from the temperature computation and/or using a spherical rather than a projected aperture has a negligible effect on the intrinsic scatter.

There is a slight drop in the intrinsic scatter between $z = 0$ and $z \approx 1$ that

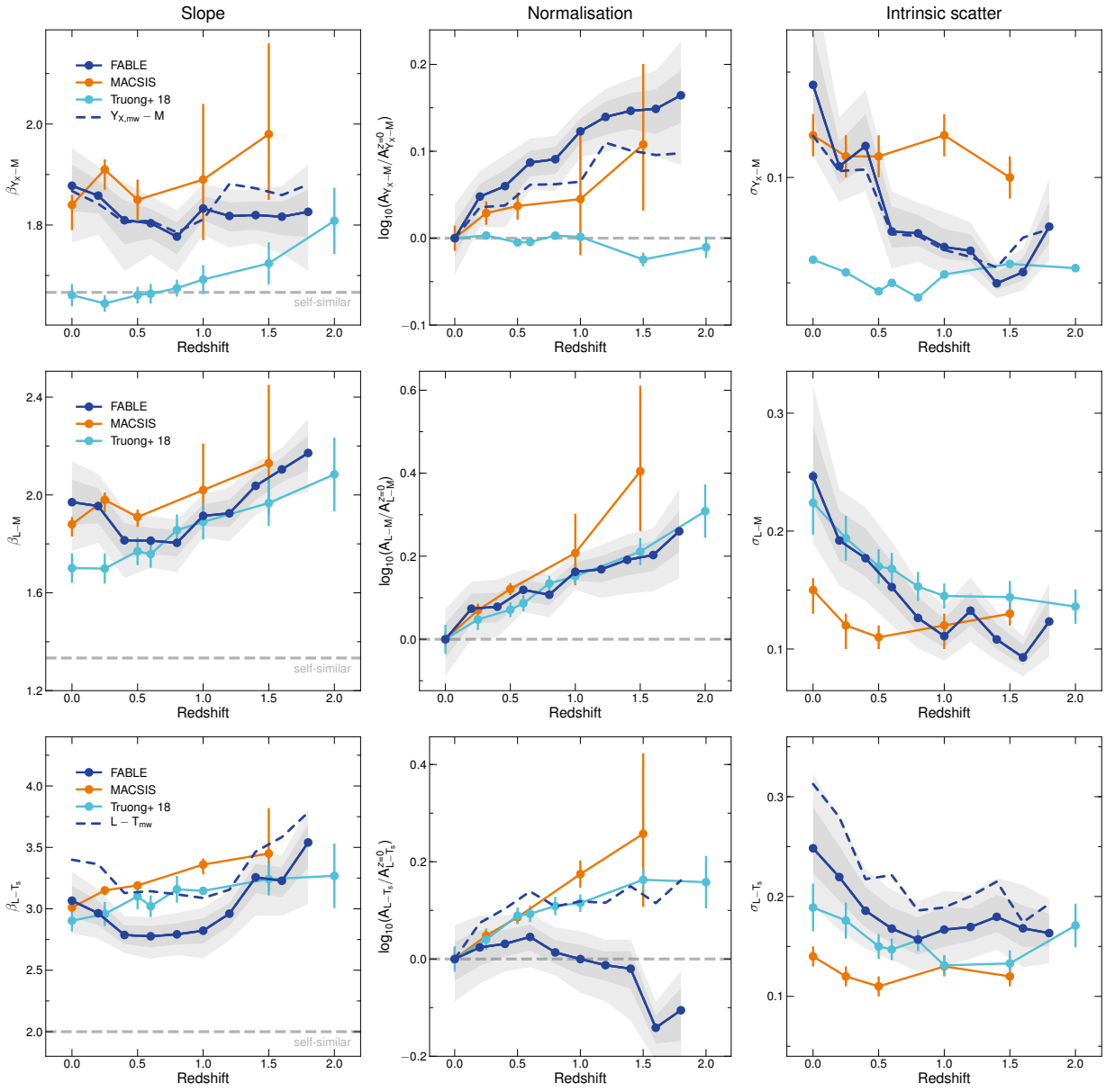


Figure 3.5: As Fig. 3.3 and 3.4 for the redshift evolution of the Y_X–total mass relation (top row), the X-ray luminosity–total mass relation (middle row) and the X-ray luminosity–temperature relation (bottom row). Blue solid and dashed lines correspond to the FABLE relations based on the spectroscopic or mass-weighted temperature, respectively. The uncertainties on the intrinsic scatter of the Y_X–total mass relation are unknown for T18.

may be driven by a decrease in the prevalence of thermal bubble injections by radio-mode AGN towards higher redshift. T18 also predict a drop in scatter with increasing redshift although, like the FABLE result, the trend is only marginally significant.

3.3.2.3 Y_X–*total mass scaling relation*

The top row of Fig. 3.5 shows the redshift evolution of the best-fitting Y_X–total mass relation, where Y_X is defined as the product of the gas mass within r_{500} and the average temperature within the core-excised aperture $(0.15\text{–}1)\,r_{500}$. We calculate Y_X using either the spectroscopic temperature (solid line) or the mass-weighted temperature (dashed line).

Y_X–total mass slope: At low redshift the Y_X–total mass relation is significantly steeper than the self-similar expectation ($\beta_{Y_X-M} = 5/3$), as has been found in several previous simulation studies (e.g. Short et al. 2010; Stanek et al. 2010; Planelles et al. 2014; Le Brun et al. 2016). At $z = 0$ for example we find a slope of $1.88^{+0.04}_{-0.05}$ for the spectroscopic temperature-based relation, which is consistent with the MACSIS result ($1.84^{+0.02}_{-0.05}$) as well as the observational constraints of Arnaud et al. (2007) (1.82 ± 0.09), Eckmiller et al. (2011) (1.82 ± 0.07) and Mahdavi et al. (2013) (1.79 ± 0.22). T18 on the other hand find a self-similar slope (1.66 ± 0.02 at $z = 0$), consistent with the simulations of Fabjan et al. (2011) and Biffi et al. (2014) and the observational findings of Lovisari et al. (2015) (1.67 ± 0.08) and Mantz et al. (2016b) (1.63 ± 0.04).

The slope of the Y_X–total mass relation is approximately equal to the product of the gas mass–total mass slope and the inverse of the total mass–temperature slope. Both components deviate from their self-similar values but in opposite directions. This is one of the motivating reasons for using the Y_X variable as a mass proxy (Kravtsov et al., 2006). For example, in T18 the steeper than self-similar gas mass–total mass and total mass–temperature slopes cancel, yielding an approximately self-similar Y_X–total mass relation at $z \lesssim 1$. In contrast, FABLE and MACSIS predict a steeper than self-similar Y_X–total mass relation at all redshifts since the gas mass–total mass relation deviates from self-similarity to a greater degree than the total mass–temperature relation. Physically, non-gravitational processes such as star formation and AGN feedback remove gas from the ICM and raise the average temperature of the hot phase (by removing cold gas via star formation and raising the entropy of the gas via thermal energy injection by AGN). Our results, and observations that measure a steeper than self-similar slope, imply that the removal of gas steepens the Y_X–total mass relation to a greater degree than the corresponding increase in the temperature of the remaining gas.

The slope is approximately independent of redshift within the uncertainties, in agreement with MACSIS. Similarly, T18 find a slope that is constant up to $z \sim 1$ although there is an increase in slope toward higher redshifts that reflects the

increase in slope of their gas mass–total mass relation (due to the shrinking mass range of their sample) and the decrease in slope of their total mass–temperature relation (due to incomplete thermalisation of gas in low-mass objects), as discussed in the previous two sections.

Y_X–total mass normalisation: A number of previous simulation studies have found that the normalisation of the $Y_X - M_{500}$ relation evolves in a self-similar manner (e.g. Kravtsov et al. 2006; Nagai et al. 2007b; Short et al. 2010; Fabjan et al. 2011), in agreement with the T18 result. In contrast, FABLE and MACSIS predict a positive, albeit mild, evolution with respect to self-similarity. This reflects the positive evolution of the gas mass–total mass normalisation that is somewhat, but not completely, offset by the mild positive evolution in the total mass–temperature relation. The evolution is slightly stronger for the spectroscopic temperature-based relation (solid line) than for the mass-weighted temperature relation (dashed line) due to redshifting of the X-ray emission, which biases the spectroscopic temperature high (see Section 3.3.2.2). Le Brun et al. (2016) also find a mild positive evolution in their low-mass sample, although for high-mass haloes ($M_{500} > 10^{14} M_\odot$) the situation is reversed. We point out that beyond self-similar evolution in the $Y_X - M_{500}$ relation would have important implications for the use of Y_X as a mass proxy, as many observational studies assume self-similar evolution when estimating total masses from this relation (e.g. Maughan et al. 2008, 2012; Bartalucci et al. 2017).

Y_X–total mass intrinsic scatter: Most previous simulation works predict a low level of intrinsic scatter in the Y_X–total mass relation at $z = 0$, for example Short et al. (2010) (0.05), Stanek et al. (2010) (0.05), Fabjan et al. (2011) (0.05) and Planelles et al. (2014) (0.08). In fact many observational studies find that it is dominated by the statistical scatter (e.g. Sun et al. 2009; Vikhlinin et al. 2009; Lovisari et al. 2015). Observations that are able to constrain the intrinsic scatter vary significantly in their measurements, for example Arnaud et al. (2007) (0.04), Mahdavi et al. (2013) (0.10 ± 0.02), Mantz et al. (2016b) (0.080 ± 0.007) and Eckmiller et al. (2011) (0.264).

The intrinsic scatter in FABLE at $z = 0$ is $0.14^{+0.02}_{-0.02}$ and $0.12^{+0.03}_{-0.02}$ for the spectroscopic and mass-weighted temperature-based relations respectively. These values are consistent with the MACSIS value (0.12 ± 0.01) but larger than most other theoretical and observational constraints, with the exception of Eckmiller et al. (2011) who measure an intrinsic scatter of 0.350 and 0.218 dex for their group and cluster samples, respectively, and 0.264 dex for their combined sample. Like Eckmiller et al. (2011) we also find a slight mass dependence in the scatter, although it is stronger for the mass-weighted temperature than the spectroscopic one. For

example, when including only cluster-scale haloes with $M_{500} > 10^{14} M_{\odot}$ in the $z = 0$ sample we find a smaller intrinsic scatter of $0.08^{+0.03}_{-0.01}$ for the mass-weighted temperature relation but a similar scatter for the spectroscopic temperature relation ($0.13^{+0.04}_{-0.02}$). This is because the latter has increased intrinsic scatter at the high mass end due to scatter in the spectroscopic temperature bias in massive clusters.

Unlike MACSIS and T18, which predict a roughly constant scatter with redshift, the intrinsic scatter decreases from $z = 0$ to $z \approx 1$. This largely reflects the redshift trend of the intrinsic scatter in the gas mass–total mass relation (Fig. 3.3). As discussed in Section 3.3.2.1, this may be caused by the increased importance of radio-mode AGN feedback at low redshift or the reduced efficiency with which AGN feedback can expel gas beyond r_{500} at high redshift.

3.3.2.4 *X-ray luminosity–total mass scaling relation*

The best-fitting parameters of the bolometric X-ray luminosity–total mass relation are shown in the middle row of Fig. 3.5.

X-ray luminosity–total mass slope: The bolometric X-ray luminosity of the ICM largely depends on the total mass of gas and its distribution, with line emission becoming dominant only in low-temperature ($T \lesssim 2$ keV) systems. As such, the evolution of the X-ray luminosity–total mass relation shares many similarities with that of the gas mass–total mass relation. Indeed, the slopes of both relations are significantly steeper than self-similar at all redshifts in FABLE, MACSIS and T18. At $z = 0$ we find a slope of $1.97^{+0.10}_{-0.08}$, consistent with the MACSIS prediction ($1.88^{+0.03}_{-0.05}$) but slightly steeper than T18 (1.70 ± 0.06).

Observational studies tend to agree on a steeper than self-similar slope, although the exact value can vary. Most observational constraints are in good agreement with the FABLE and MACSIS predictions, for example, Maughan (2007) (1.96 ± 0.10), Pratt et al. (2009) (1.96 ± 0.11 or 2.08 ± 0.13 after correcting for Malmquist bias) and Giles et al. (2017) (2.22 ± 0.24). Conversely, Mantz et al. (2010) and Reichert et al. (2011) find considerably shallower slopes of 1.63 ± 0.06 and 1.52 ± 0.09, respectively, although the former study uses a particularly massive cluster sample while the latter may be biased by the inhomogeneity of their dataset, which is drawn from multiple sources in the literature.

We point out that, whilst our X-ray luminosities are measured within a projected aperture of radius r_{500}, those of MACSIS and T18 are measured within a spherical, core-excised aperture of radius $(0.15\text{–}1)\, r_{500}$. A projected aperture mimics observations but can cause the X-ray luminosity to be boosted by hot gas along the line of sight. This slightly lowers the slope of the luminosity–based relations

because projection has a larger impact on group-scale objects than on massive clusters since the former are more affected by contamination with the latter. In the general field, significant overlap between haloes is fairly rare. However, for the secondary haloes in our zoom-in simulations this occurs more frequently, which leads to a non-negligible effect on the X-ray luminosity–total mass slope. Indeed, the increase in slope found by switching from a projected to a spherical aperture of the same radius is small compared with the uncertainties (e.g. from $1.97^{+0.10}_{-0.08}$ to $2.02^{+0.11}_{-0.09}$ at $z = 0$) but is systematic across all redshift bins. Excluding the core ($< 0.15\,r_{500}$) from the spherically integrated X-ray luminosity leads to overall slightly shallower slopes, as was also found in T18. Coincidentally, this change is approximately equal and opposite to the effect of projection so that the slope and its evolution are unchanged by switching from the projected aperture used in Fig. 3.5 to a spherical, core-excised aperture as in MACSIS and T18.

The slope of the X-ray luminosity–total mass relation shows some redshift dependence. At $z \lesssim 1$ this reflects the change in slope of the gas mass–total mass relation (Section 3.3.2.1). At $z \gtrsim 1$ the slope of the X-ray luminosity–total mass relation increases with increasing redshift similarly to MACSIS. Like B17 we attribute this redshift evolution to the combined redshift trends of the gas mass–total mass and total mass–temperature slopes, which are negligible on their own but combine into a mild redshift evolution in the X-ray luminosity–total mass slope. The FABLE and MACSIS total mass–temperature slopes show opposing redshift trends at $z \gtrsim 1$, however, the FABLE sample is significantly less massive than MACSIS, particularly at high redshift. This modifies the temperature dependence of the X-ray luminosity due to the increasing contribution of line emission at lower temperatures (Maughan, 2013). Indeed, with a more massive sample of haloes with $M_{500} > 10^{14} E(z)^{0.5} M_{\odot}$ we find no significant redshift evolution in the X-ray luminosity–total mass slope.

X-ray luminosity–total mass normalisation: The normalisation of the X-ray luminosity–total mass relation has a positive redshift evolution, being ~ 60 per cent higher at $z = 1.5$ than at $z = 0$ relative to self-similarity. MACSIS and T18 predict a very similar increase in normalisation within the uncertainties. As for the gas mass–total mass relation, this evolution is largely due to the increased binding energy at fixed total mass with increasing redshift, which inhibits the reduction in the gas mass (and thus X-ray luminosity) via AGN feedback.

When measuring the X-ray luminosity within a spherical, core-excised aperture, the normalisation evolves slightly faster (e.g. ~ 75 per cent between $z = 0$ and $z = 1.5$). The difference is almost entirely due to projection effects rather than

excision of the cluster core. At low redshift the X-ray luminosity at fixed total mass is biased high by hot gas along the line of sight that lies predominantly in hotter, more massive objects. At higher redshift the projection bias at fixed mass is smaller because there are fewer objects above that mass. Hence, the normalisation evolves less rapidly when the luminosity is projected rather than spherically integrated. Note however that we are likely overestimating the effect of projection due to our inclusion of secondary haloes from the zoom-in simulations, which are more likely to overlap with a more massive halo, and because we do not perform background subtraction in our X-ray analysis as would be carried out in real observations.

We caution that, because the slope of the X-ray luminosity–total mass relation varies substantially with redshift, the evolution of the normalisation is sensitive to the choice of pivot point. In particular, choosing a larger (smaller) mass for the pivot point makes the apparent evolution in the normalisation more (less) positive. For example, at pivot points of $10^{14} M_{\odot}$ and $5 \times 10^{14} M_{\odot}$ keV the increase in normalisation between $z = 0$ and $z = 1.5$ relative to self-similarity is ~ 45 and ~ 70 per cent, respectively. A similar change is found for the other simulation relations as they evolve similarly in slope.

X-ray luminosity–total mass intrinsic scatter: The intrinsic scatter of the X-ray luminosity–total mass relation is $0.25^{+0.04}_{-0.03}$ at $z = 0$, which agrees with the T18 value (0.22 ± 0.03) but is somewhat higher than MACSIS ($0.15^{+0.01}_{-0.02}$). Observational constraints on the intrinsic scatter span a similar range of values, for example Sun (2012) and Giles et al. (2017) measure 0.25 ± 0.05 and 0.30 ± 0.05, respectively, while Maughan (2007), Pratt et al. (2009) and Mantz et al. (2010) derive smaller values of 0.17 ± 0.02, 0.166 ± 0.026 and 0.185 ± 0.019, respectively.

Much of the variation between studies can be attributed to sample selection. For example, lower mass haloes exhibit somewhat larger intrinsic scatter so that restricting our sample to haloes with $M_{500} > 10^{14} E(z)^{-0.5} M_{\odot}$ slightly reduces the scatter at all redshifts (e.g. from $0.25^{+0.04}_{-0.03}$ to $0.21^{+0.04}_{-0.02}$ at $z = 0$). Furthermore, several studies have found that the scatter in X-ray luminosity at fixed mass or temperature is dominated by the cluster core regions, particularly for mixed samples of relaxed, cool core clusters and unrelaxed, morphologically disturbed ones (e.g. Markevitch 1998; Maughan 2007; Pratt et al. 2009; Maughan et al. 2012). Indeed, when excising the core from the projected aperture we find that the scatter at $z = 0$ decreases slightly from $0.25^{+0.04}_{-0.03}$ to $0.22^{+0.04}_{-0.03}$, with a similar decrease at higher redshifts. This is a smaller effect than is typically found in observations however, which may be due to a lack of strong cool core clusters in our sample.

At $z \lesssim 0.5$ all three simulations predict a drop in the intrinsic scatter with

increasing redshift. This is in qualitative agreement with Mantz et al. (2016b) who find tentative evidence for evolution in the intrinsic scatter of X-ray luminosity at fixed mass for a large sample of clusters at $z \lesssim 0.5$. At $0.5 \lesssim z \lesssim 1$ the intrinsic scatter in FABLE and T18 continues to fall with increasing redshift. This is consistent with Maughan (2007) who measure a significantly larger intrinsic scatter for a subset of clusters at $0.1 < z < 0.5$ compared with a subset at $0.5 < z < 1.3$ (approximately a factor of two increase). At $z \gtrsim 1$ the simulations converge on a roughly redshift-independent scatter of 0.10–0.15 dex.

3.3.2.5 *X-ray luminosity–temperature scaling relation*

The bottom row of Fig. 3.5 shows the best-fitting parameters of the X-ray luminosity–temperature relation. Solid and dashed lines correspond to the relation based on the spectroscopic and mass-weighted temperature, respectively.

X-ray luminosity–temperature slope: The slope of the X-ray luminosity–temperature relation is steeper than the self-similar expectation ($\beta_{L-T} = 2$) at all redshifts in the FABLE, MACSIS and T18 simulations. At $z = 0$ the slope of the FABLE relation is $3.07^{+0.11}_{-0.15}$ and $3.40^{+0.19}_{-0.20}$ for the spectroscopic and mass-weighted temperature-based relations, respectively. These values are slightly higher than T18 (2.903 ± 0.086) but statistically consistent with MACSIS (3.07 ± 0.04) and the majority of observational constraints, for example Pratt et al. (2009) (3.35 ± 0.32), Hilton et al. (2012) (3.04 ± 0.16), Sun (2012) (3.03 ± 0.01), Giles et al. (2016) (3.08 ± 0.15) and Zou et al. (2016) (3.28 ± 0.33).

The slope of the X-ray luminosity–temperature relation is approximately equal to the product of the X-ray luminosity–total mass and total mass–temperature slopes. Since the latter shows negligible redshift evolution, the X-ray luminosity–temperature slope increases mildly with redshift as for the X-ray luminosity–total mass relation. The slope of the spectroscopic temperature-based relation is biased low compared with the mass-weighted temperature relation due to the mass dependence of the spectroscopic temperature bias described in Section 3.3.2.2.

The slope increases slightly when measuring the luminosity and temperature within a spherical rather than a projected aperture (from $3.07^{+0.11}_{-0.15}$ to $3.23^{+0.15}_{-0.15}$ at $z = 0$ and similarly at higher redshift). This is because gas that overlaps with the cluster in projection causes the luminosity and spectroscopic temperature to be biased high and low, respectively, particularly in low-mass objects. Excising the cluster core has a negligible effect on the X-ray luminosity–temperature slope as the X-ray luminosity–total mass and total mass–temperature slopes show opposing behaviour.

X-ray luminosity–temperature normalisation: The X-ray luminosity at fixed mass-weighted temperature evolves positively with respect to self-similarity, with clusters at the pivot temperature (3 keV) having ~ 35 per cent higher luminosity at $z = 1.5$ compared with $z = 0$. This represents the combined evolution of the X-ray luminosity–total mass and total mass–temperature relations. In contrast, the spectroscopic temperature-based relation shows negligible evolution. A similar offset was found for the total mass–temperature relation (Section 3.3.2.2) and is due to an increasing spectroscopic temperature bias caused by redshifting of the low-energy X-ray emission beyond the X-ray bandpass. As for the X-ray luminosity–total mass relation, the evolution in the normalisation is somewhat dependent on the pivot point. In particular, using a higher (lower) pivot point causes the normalisation to evolve more (less) positively in each case. For example, at a pivot point of 5 keV the increase in X-ray luminosity between $z = 0$ and $z = 1.5$ rises to ~ 45 per cent.

The positive evolution in normalisation predicted by MACSIS, T18 and the mass-weighted temperature relation of FABLE agrees with the results of Giles et al. (2016) for the XXL-100-GC sample, which spans a redshift range of $0.05 < z < 1.05$. Assuming a redshift-independent slope, Giles et al. (2016) find that the normalisation of the X-ray luminosity–temperature relation evolves as $E(z)^{1.64\pm0.77}$. This corresponds to an increase in normalisation of ≈ 0.15 dex between $z = 0$ and $z = 1$ relative to self-similarity, which is in good agreement with the simulation predictions, albeit with large uncertainties. In contrast, the observational studies by Reichert et al. (2011), Hilton et al. (2012) and Clerc et al. (2014) measure a negative evolution in the normalisation of the X-ray luminosity–temperature relation up to $z \sim 1.5$. The Reichert et al. (2011) result may be affected by their use of an externally-calibrated local baseline relation for their comparison to high-redshift clusters. Furthermore, Reichert et al. (2011) measure a slightly steeper slope for their $z > 0.8$ clusters compared with their local relation, which may drive some of the evolution in the normalisation. Hilton et al. (2012) find no evidence for evolution in the slope of the relation since $z \sim 1.5$, however they do not account for selection biases as in the other studies. Lastly, Giles et al. (2016) attribute the difference compared to Clerc et al. (2014) largely to the choice of local baseline relation for which differences of less than 3-sigma can change the inferred evolution from positive to negative.

It is worth pointing out that the positive evolution indicated in Fig. 3.5 represents an even greater departure from self-similarity than it appears. This is because the factor $E(z)$ that we have incorporated into the normalisation of the X-ray luminosity–temperature relation in order to factor out the expected self-similar evolution is no longer the appropriate expected scaling given that the slope of the

X-ray luminosity–total mass relation has been shown to depart from self-similarity. If we assume that the X-ray luminosity–total mass and total mass–temperature relations take the form of equation 3.1 with slopes $\beta_{\rm LM}$ and $\beta_{\rm MT}$ and $E(z)$ exponents $\gamma_{\rm LM}$ and $\gamma_{\rm MT}$, respectively, then these equations can be combined to show that the luminosity, L, scales with temperature, T, as

$$L \propto E(z)^{(\gamma_{\rm LM} + \gamma_{\rm MT}\beta_{\rm LM})}\ T^{\beta_{\rm LM}\beta_{\rm MT}}. \tag{3.3}$$

From this equation it is clear that, even if the X-ray luminosity–total mass and total mass–temperature relations evolve self-similarly (i.e. $\gamma_{\rm LM}$ and $\gamma_{\rm MT}$ equal their self-similar values), any departure of the slope $\beta_{\rm LM}$ from the self-similar value will alter the expected evolution of the X-ray luminosity–temperature relation. As $\beta_{\rm LM}$ is steeper than the self-similar value (Section 3.3.2.4) and $\gamma_{\rm MT} < 0$ (both in the self-similar scenario and in our simulations), self-similar evolution would imply that the normalisation of the X-ray luminosity–temperature relation should evolve less rapidly than $E(z)$. For example, using the $z = 0$ value for $\beta_{\rm LM}$ and the self-similar values $\gamma_{\rm LM} = 7/3$ and $\gamma_{\rm MT} = -1$ yields a normalisation that evolves approximately as $E(z)^{0.4}$. As a result, the positive evolution of the normalisation shown in Fig. 3.5 somewhat underestimates the departure from the expected evolution. Similar reasoning applies to the MACSIS and T18 relations, which also predict a steeper than self-similar X-ray luminosity–total mass slope.

X-ray luminosity–temperature intrinsic scatter: The intrinsic scatter about the X-ray luminosity–temperature relation is somewhat higher than the MACSIS and T18 predictions due to the increased scatter of lower mass haloes. When restricting our $z = 0$ sample to haloes with $M_{500} > 10^{14} M_\odot$ (the same as MACSIS and T18 at $z = 0$) the intrinsic scatter drops from $0.25^{+0.04}_{-0.03}$ to $0.18^{+0.05}_{-0.02}$ for the spectroscopic temperature relation, which lies in between the $z = 0$ values for MACSIS (0.14 ± 0.01) and T18 (0.19 ± 0.02). These values are somewhat smaller than the observational constraints of Pratt et al. (2009) (0.32 ± 0.058), Hilton et al. (2012) (0.27 ± 0.03), Maughan et al. (2012) (0.29 ± 0.02) and Sun (2012) (0.24 ± 0.01) but consistent with recent observational constraints from Giles et al. (2016) (0.20 ± 0.03) and Zou et al. (2016) (0.20 ± 0.05).

We find that the intrinsic scatter drops slightly with increasing redshift, falling from $0.25^{+0.04}_{-0.03}$ at $z = 0$ to $0.17^{+0.02}_{-0.02}$ at $z = 1$ and remaining roughly constant at higher redshift. This is in good agreement with the results of Hilton et al. (2012) who measure the evolution of the X-ray luminosity–temperature relation with 211 clusters up to $z \sim 1.5$ and find a similar decrease in the intrinsic scatter with redshift, falling from 0.33 ± 0.04 dex at $z < 0.25$ to 0.24 ± 0.05 at $0.5 < z < 1.5$.

3.4 SZ–TOTAL MASS RELATION

A complementary means of measuring the thermodynamic properties of gas in groups and clusters is via the thermal SZ effect, which arises from the inverse Compton scattering of CMB photons from energetic electrons in the hot ICM (see Section 1.5). The SZ signal, $Y_{\rm SZ}$, is characterised by the integrated Compton y parameter as defined in equation 2.1.

In this section we explore the scaling between SZ flux, $Y_{\rm SZ}$, and halo mass, M_{500}. We fit the $Y_{\rm SZ} - M_{500}$ relations with a power law as described in Section 3.2.3 to the sample defined in Section 3.2.4. In Section 3.4.1 we compare the FABLE relation to observed clusters from the *Planck* and SPT-SZ catalogues out to $z \sim 0.8$ and $z \sim 1.2$, respectively, measuring $Y_{\rm SZ}$ within an aperture appropriate for the comparison. In Section 3.4.2 we study the redshift evolution of the $Y_{500} - M_{500}$ relation ($Y_{\rm SZ}$ measured within a radius r_{500}) and in Section 3.4.3 we investigate how different predictions for the $Y_{500} - M_{500}$ relation and its redshift evolution affect the expected number of clusters in an SZ-selected survey such as SPT-3G.

3.4.1 *Comparison with Planck and SPT*

The second *Planck* SZ source catalogue (Planck Collaboration XXVII, 2016) is the largest SZ-selected sample of galaxy clusters to-date, containing over one thousand confirmed clusters out to $z \sim 1$. The catalogue contains estimates for the integrated flux of each cluster measured within a circular aperture of radius $5\,r_{500}$, which we denote $\mathrm{Y}_{5\mathrm{r}_{500}}$. Total cluster masses, M_{500}, are estimated assuming a scaling relation between the SZ flux and M_{500} calibrated on X-ray hydrostatic masses (Planck Collaboration XX, 2014) given the *Planck* posterior information on the size-flux correlation. These mass estimates will depend on the assumed scaling relation, although Planck Collaboration XXVII (2016) show that they agree with external X-ray and optical data with low scatter. For the simulated objects we calculate fluxes within a spherical aperture of radius $5\,r_{500}$. Using a spherical rather than a circular aperture has a negligible effect on $\mathrm{Y}_{5\mathrm{r}_{500}}$ over the range of halo masses in common with the *Planck* clusters ($\gtrsim 10^{14} M_{\odot}$) and thus does not affect our comparison. However, for lower mass systems the SZ flux measured within a circular aperture can be boosted by hot gas that overlaps in projection and for this reason we use a spherical aperture to avoid biasing the best-fitting $\mathrm{Y}_{5\mathrm{r}_{500}} - \mathrm{M}_{500}$ relation, which is fit to $M_{500} \geq 3 \times 10^{13}\, E(z)^{-0.5}\, M_{\odot}$ (see Section 3.2.4).

In Figure 3.6 we plot $\mathrm{Y}_{5\mathrm{r}_{500}}$ as a function of M_{500} at five redshifts between $z = 0$ and $z = 0.8$. We scale $\mathrm{Y}_{5\mathrm{r}_{500}}$ by the square of the angular diameter distance of each

cluster and multiply by the self-similar scaling factor $E(z)^{-2/3}$. At each redshift we select a comparison sample of *Planck* clusters with a similar median redshift. The redshift range that defines each sample is given in the legend. Due to the large number of sources, we bin the *Planck* clusters in halo mass bins of width 0.2 dex and plot the median and 1-sigma intrinsic scatter. For bins containing fewer than ten objects we plot the individual clusters. In addition, at $z = 0$ we compare with the SZ flux measurements of Planck Collaboration XI (2013) for a large sample of locally brightest galaxies (LBGs). These galaxies were selected to be predominantly central galaxies and therefore form a natural extension of the cluster relation to lower masses. Planck Collaboration XI (2013) bin the LBGs by stellar mass and then estimate an "effective" halo mass for each bin using the semi-analytic galaxy formation simulation of Guo et al. (2011). Since this procedure has a non-negligible dependence on the semi-analytic model (Wang et al., 2016), we compare to the Planck Collaboration XI (2013) relation recalibrated by Wang et al. (2016) using weak lensing masses. We convert the Y_{500} values reported in Planck Collaboration XI (2013) back into the measured flux, $\mathrm{Y}_{5\mathrm{r}_{500}}$, by multiplying by the factor 1.796, which is the conversion factor corresponding to the spatial template used in their matched filter, the universal pressure profile (Arnaud et al., 2010).

At $z = 0$ the FABLE relation is in excellent agreement with the Wang et al. (2016) LBG relation over their full mass range ($3 \times 10^{13} M_\odot \lesssim M_{500} \lesssim 3 \times 10^{14} M_\odot$). In comparison, the *Planck* clusters have slightly higher SZ flux at fixed halo mass. The most likely explanation for this offset is an X-ray hydrostatic mass bias of ~ 30 per cent, which would bias the *Planck* cluster relation relative to the weak lensing-calibrated LBG relation and the simulations. This is consistent with the ~ 30 per cent X-ray mass bias required to explain the offset in the total mass–spectroscopic temperature relation between FABLE and observational constraints based on X-ray hydrostatic masses (see e.g. Fig. 3.1 and 3.2). A similar offset between the *Planck* clusters and the FABLE relation persists to higher redshift, although at $z \geq 0.6$ we rely on the extrapolation of the best-fitting relation as there are few, if any, FABLE systems as massive as the *Planck* clusters. This suggests that the X-ray mass bias remains at a similar level with increasing redshift, as found for example in Nagai et al. (2007a), Henson et al. (2017) and Hurier & Angulo (2018).

The best-fitting power law relation is a good description of the $\mathrm{Y}_{5\mathrm{r}_{500}} - M_{500}$ scaling over a wide range of masses, from massive clusters ($\sim 10^{15} M_\odot$) to low mass galaxy groups ($\sim 10^{13} M_\odot$). The slope of the relation is consistent with the self-similar prediction at all redshifts. We also find no evidence for a systematic redshift trend in the normalisation or intrinsic scatter. This is understandable as the SZ flux is integrated out to several times the virial radius and is thus fairly

insensitive to non-gravitational processes such as AGN feedback.

In Figure 3.7 we perform a similar comparison with galaxy clusters discovered in the 2500 deg^2 SPT-SZ survey (Bleem et al., 2015) at $0.2 \leq z \leq 1.2$. We use the cluster catalogue described in Bocquet et al. (2018), which uses an updated calibration of the photometric redshifts from Bleem et al. (2015) plus additional spectroscopic redshifts. Bocquet et al. (2018) also perform a weak-lensing calibrated cosmological analysis of the SPT data and derive cluster mass estimates marginalized over the uncertainties in the scaling relation and the cosmological parameters. We use these mass estimates in our comparison but note that our conclusions are unchanged if instead we use the fiducial mass estimates presented in Bleem et al. (2015). The SPT SZ flux measurements are integrated within a circular aperture of radius 0.75 arcminutes. We measure the SZ flux within the same radius, integrating the entire simulation volume along the line-of-sight.

The FABLE clusters are a good match to the SPT data at all redshifts. The best-fitting relation tends to overestimate the SZ flux of the most massive SPT clusters, however this is because the fit is biased towards the lowest mass systems, which prefer a steeper slope. Although we do not have a large number of haloes in a common mass range with the SPT clusters, those which do overlap are in excellent agreement. The handful of outliers with high SZ flux at the low mass end of the relation are objects that overlap with more massive haloes in projection, which boosts their apparent flux. Note, however, that these are mostly secondary haloes in the zoom-in simulations for which the probability of such an alignment is significantly increased.

Whilst the relatively large *Planck* aperture integrates the SZ flux to the cluster outskirts, SPT is sensitive to the cluster core regions. For example, at $z = 0.2$ the 0.75 arcminute aperture radius corresponds to a physical radius of only 153 kpc. For a typical mass of $M_{500} \sim 5 \times 10^{14} M_{\odot}$ with $r_{500} \sim 1$ Mpc this corresponds to roughly $0.15\, r_{500}$, which characterises the cluster core. Therefore, our agreement with the SPT and *Planck* clusters suggests that the SZ flux is realistically distributed from the core to the cluster outskirts. Indeed, we showed in Section 2.6.4 that the pressure profiles of FABLE clusters are in good agreement with observations of local clusters from the REXCESS sample (Böhringer et al., 2007; Pratt et al., 2010).

3.4.2 *The $Y_{500} - M_{500}$ relation*

In this section we investigate the $Y_{500} - M_{500}$ relation, where Y_{500} is the SZ flux measured within a spherical aperture of radius r_{500}. In Fig. 3.8 we show the $Y_{500} - M_{500}$ relation at $z = 0$ and in Fig. 3.9 we plot the parameters of the best-

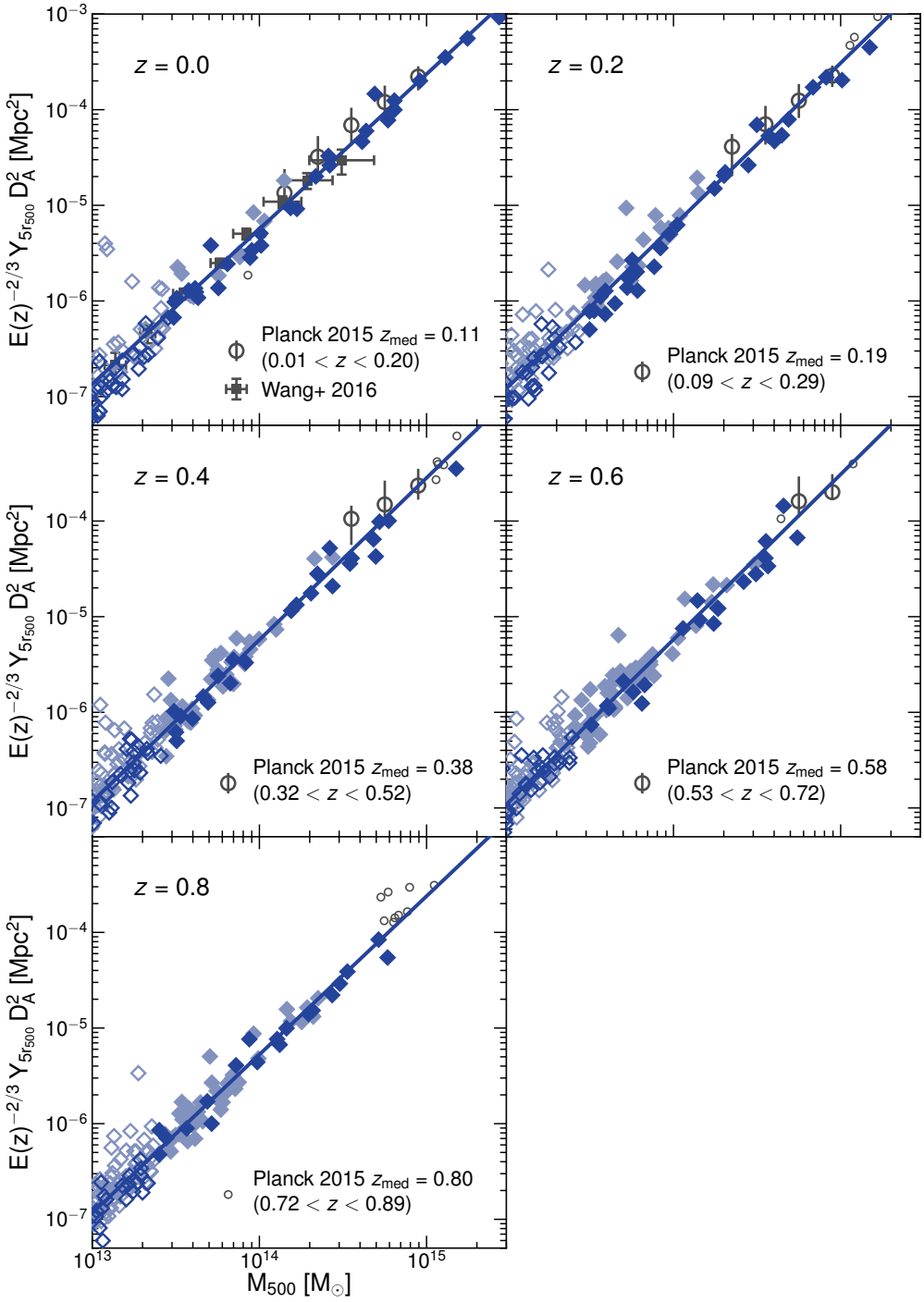


Figure 3.6: The SZ signal, $Y_{5r_{500}}$, as a function of total mass for FABLE haloes (diamonds; styles as in Fig. 3.1) at various redshifts compared to *Planck* data (grey circles). $Y_{5r_{500}}$ is the Compton y parameter integrated within a spherical aperture of radius $5\,r_{500}$. We scale $Y_{5r_{500}}$ by the square of the angular diameter distance and multiply by $E(z)^{-2/3}$ to factor out self-similar redshift evolution. The solid line shows the best-fitting power-law relation to haloes of mass $M_{500} > 3 \times 10^{13} M_{\odot}$ (filled diamonds). Open circles with error bars show the median relation and 1-sigma intrinsic scatter in total mass bins of width 0.2 dex for *Planck* clusters in the redshift range indicated in the panel legends. For bins containing fewer than ten objects we plot individual clusters as small open circles.

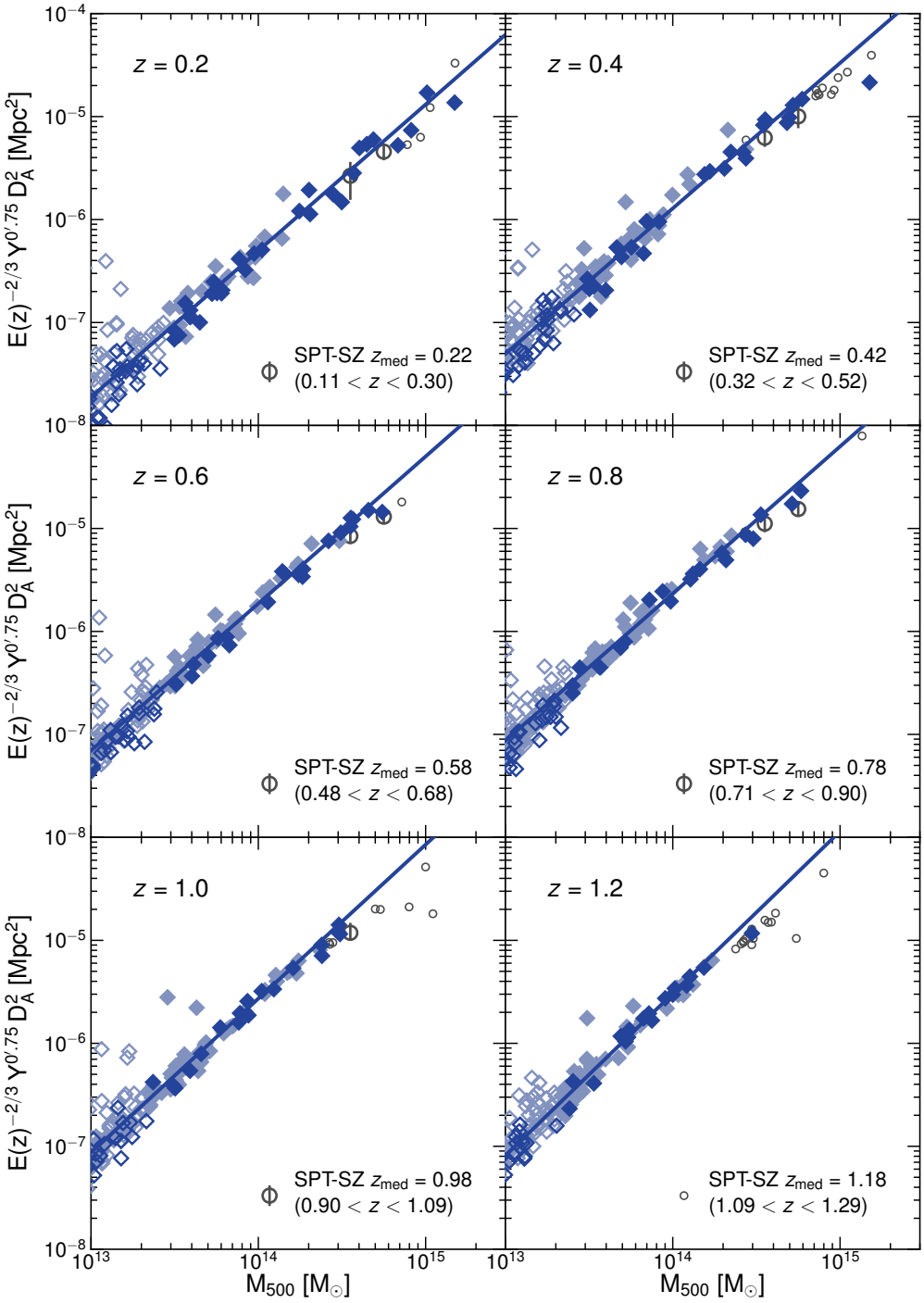


Figure 3.7: The SZ signal integrated within a circular aperture of radius 0.75 arcminutes as a function of total mass at different redshifts compared with clusters detected in the 2500 deg^2 SPT-SZ survey (Bleem et al., 2015; Bocquet et al., 2018). Symbol styles are similar to those in Fig. 3.6.

fitting relation as a function of redshift. We compare to the best-fitting relations from the MACSIS and Planelles et al. (2017) simulations as well as observationally derived relations from Andersson et al. (2011), Planck Collaboration XX (2014) and Nagarajan et al. (2018).

The simulations reported in Planelles et al. (2017) were analysed in terms of their X-ray properties in T18, which we have compared with the FABLE and MACSIS results in Section 3.3. Note that Planelles et al. (2017) use a lower mass sample than T18, consisting of ~ 100 clusters and groups with $M_{500} \gtrsim 4.1 \times 10^{13} M_{\odot}$ at $z = 0$. Andersson et al. (2011) derive the $Y_{500} - M_{500}$ relation for a sample of 15 SPT clusters observed with *Chandra* and *XMM-Newton*. Cluster candidates with the highest signal-to-noise ratio were chosen from 178 deg^2 of the sky surveyed by the SPT in 2008 (Vanderlinde et al., 2010). Total masses are estimated from the $M_{500} - Y_{\rm X}$ relation of Vikhlinin et al. (2009), which is calibrated on X-ray hydrostatic mass estimates. We note that $Y_{\rm X}$ is related but not equal to Y_{500} because the two quantities rely on differently weighted gas temperatures (X-ray spectroscopic and mass-weighted, respectively). Indeed, the $Y_{500} - Y_{\rm X}$ relation derived in Andersson et al. (2011) has an intrinsic scatter of 0.09 ± 0.04 dex and its normalisation implies an average $Y_{500}/Y_{\rm X}$ ratio of 0.82 ± 0.07. Planck Collaboration XX (2014) derive a baseline $Y_{500} - M_{500}$ relation for use in their cosmological analysis of SZ cluster counts. They take 71 clusters from the *Planck* cosmological sample with *XMM-Newton* observations and estimate their masses from the relation between Y_X and the X-ray hydrostatic mass M_{500} established for 20 local relaxed clusters by Arnaud et al. (2010). The $Y_{500} - Y_{\rm X}$ relation for this sample has an intrinsic scatter of 0.07 ± 0.01 dex and an average $Y_{500}/Y_{\rm X}$ ratio of 0.94 ± 0.02 (Planck Collaboration XXIX, 2014). Finally, Nagarajan et al. (2018) calibrate the $Y_{500} - M_{500}$ relation with SZ measurements from the APEX-SZ experiment and total cluster mass estimates derived from weak lensing. The sample consists of 27 clusters selected from the ROSAT All-Sky Survey.

3.4.2.1 *Comparison at $z \approx 0$*

To study the reliability of our model at the low mass end of the $Y_{500} - M_{500}$ scaling relation we compare to data from Wang et al. (2016) and Lim et al. (2018). We compare to the Wang et al. (2016) weak lensing-calibrated relation of the LBG sample using the Y_{500} values given in Planck Collaboration XI (2013). Lim et al. (2018) use the matched filter approach applied to the *Planck* all-sky maps to measure the SZ effect produced by local galaxy groups. The sample constitutes galaxy groups identified with a halo-based group finder from several large surveys with halo masses assigned based on the stellar mass of member galaxies. The

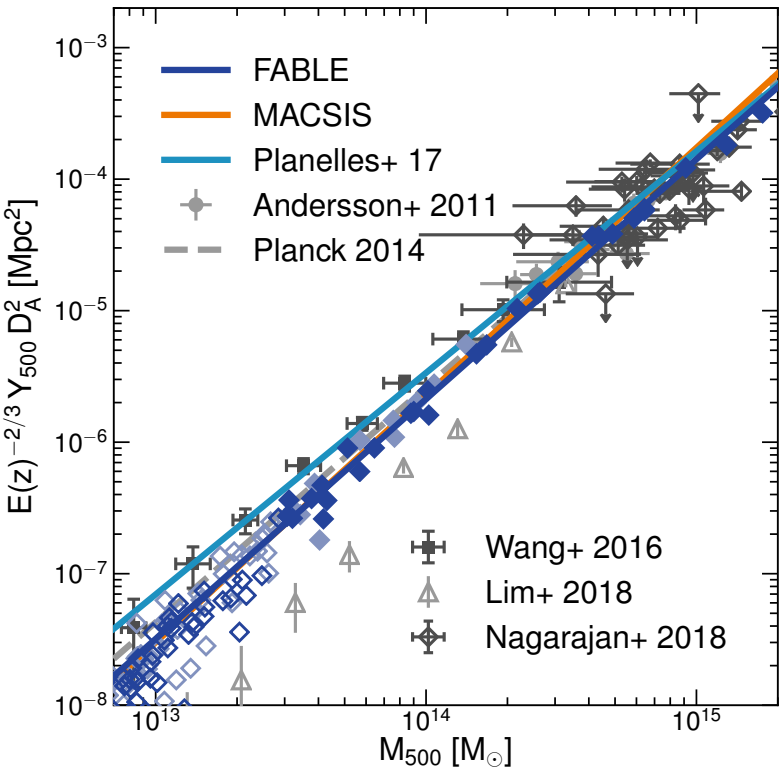


Figure 3.8: The SZ signal integrated within a spherical aperture of radius r_{500} as a function of total mass at $z = 0$ (diamonds; styles as in Fig. 3.1). Solid lines indicate the best-fitting relations for FABLE (dark blue), MACSIS (orange) and the Planelles et al. (2017) simulations (light blue). Grey error bars show data from Andersson et al. (2011), Wang et al. (2016), Lim et al. (2018) and Nagarajan et al. (2018). The grey dashed line is the baseline relation of Planck Collaboration XX (2014).

groups are stacked via their halo mass and the mean SZ flux calculated for each bin.

At $z = 0$ (Fig. 3.8) the predicted and observed best-fitting relations are in excellent agreement in the cluster regime, overlapping to within ~ 0.1 dex at $\sim 5 \times 10^{14} M_\odot$. This supports previous results that show the $Y_{500} - M_{500}$ relation is relatively insensitive to the precise implementation of non-gravitational physics in simulations (e.g. Battaglia et al. 2012; Le Brun et al. 2014; Pike et al. 2014).

On the other hand, there is a significant dispersion between relations in the regime of low-mass clusters and galaxy groups ($M_{500} \lesssim 3 \times 10^{14} M_\odot$). Andersson et al. (2011), Planck Collaboration XX (2014) and Nagarajan et al. (2018), and indeed most other observational studies, fit the SZ flux–mass relation only on cluster scales and so it is not surprising that the best-fitting relations of these studies do not fully agree at these relatively low masses. However, observational constraints at the low mass end by Wang et al. (2016) and Lim et al. (2018) present

an equally large discrepancy (compare squares and diamonds in Fig. 3.8).

The Wang et al. (2016) data lie slightly above the FABLE $Y_{500}-M_{500}$ relation, even though we have good agreement with their $Y_{5r500}-M_{500}$ relation (Figure 3.6). This suggests that there is a bias in the method used in Planck Collaboration XI (2013) to infer Y_{500} from the measured flux, Y_{5r500}. Indeed, Le Brun et al. (2015) generate mock synthetic thermal SZ maps from the cosmo-OWLS suite of simulations to show that Y_{500} estimated using the Planck Collaboration XI (2013) spatial template is biased high, for instance by a factor of two at $M_{500} \sim 3 \times 10^{13} M_{\odot}$. In fact the bias found in Le Brun et al. (2015) accounts for almost all of the offset between the Wang et al. (2016) and FABLE relations, and similarly for the extrapolated MACSIS relation.

Lim et al. (2018) obtain a significantly lower amplitude for galaxy groups compared with the Wang et al. (2016) weak lensing calibrated LBG relation and the original Planck Collaboration XI (2013) LBG relation. Lim et al. (2018) discuss two possible reasons for this difference. First is that they match all groups in the sample during the matched filtering process simultaneously, which accounts for projection by larger haloes along the line-of-sight, while Planck Collaboration XI (2013) matches individual filters separately. Second, Planck Collaboration XI (2013) bin haloes by the stellar mass of the central galaxy, which may mix haloes of different masses. On the other hand, FABLE, MACSIS and Planelles et al. (2017) all predict a significantly higher normalisation than Lim et al. (2018) that is closer to the Wang et al. (2016) relation. Furthermore, the Lim et al. (2018) relation is significantly steeper than the constraints at cluster scales, which implies a break in the $Y_{500} - M_{500}$ relation at $M_{500} \sim 3 \times 10^{14} M_{\odot}$. Such a change in slope is not apparent in the simulation relations shown here, although there is some evidence of a break in the cosmo-OWLS simulations (see fig. 2 of Lim et al. 2018).

3.4.2.2 *Evolution of the $Y_{500} - M_{500}$ relation*

Figure 3.9 shows the redshift evolution of the slope, normalisation and intrinsic scatter of the best-fitting $Y_{500} - M_{500}$ relation. For comparison we plot the observational constraints on these parameters at the median redshift of their respective samples.

$Y_{500} - M_{500}$ slope: The slope of the FABLE relation is approximately constant with redshift at $\beta_{Y_{500}-M} \sim 1.8$, which is slightly steeper than self-similar ($\beta_{Y_{500}-M} = 5/3$). This value is in good agreement with Planck Collaboration XX (2014) but slightly high compared with Andersson et al. (2011), Wang et al. (2016) and Nagarajan et al. (2018). On the other hand, limiting our sample to massive objects

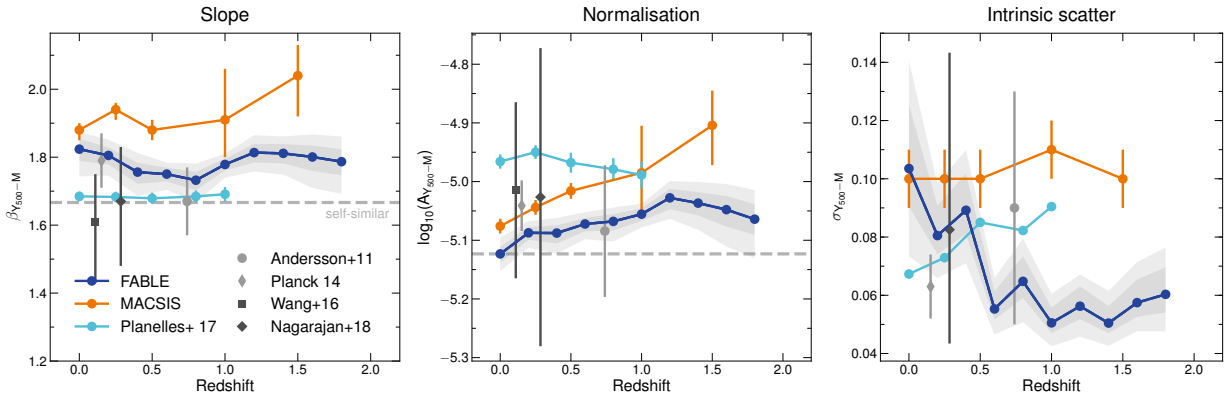


Figure 3.9: The redshift evolution of the $Y_{500} - M_{500}$ relation, where Y_{500} is the SZ signal integrated within a spherical aperture of radius r_{500}. Panels show the slope, normalisation and intrinsic scatter of the best-fitting power-law relation from left to right, respectively. Symbol styles are the same as Fig. 3.3. The light blue curves show the results of Planelles et al. (2017), who analyse the same set of simulations as T18 but using a different sample selection as described in the text. Observational constraints on the parameters are plotted as grey error bars at the median redshift of their cluster sample. Note that for Planck Collaboration XX (2014) the error bar on the normalisation indicates the statistical uncertainty and does not include the uncertainty in their X-ray mass bias parameter.

($M_{500} > 10^{14} E(z)^{-0.5} M_\odot$) yields to a shallower slope that is consistent with most observational constraints and the self-similar prediction within the uncertainties. Planelles et al. (2017) predict a close to self-similar slope out to $z = 1$, whereas the MACSIS relation is significantly steeper than self-similar out to $z = 1.5$. This likely reflects the difference in the slope of their gas mass–total mass relations (see Section 3.3.2.1). The lack of redshift evolution in the slope is consistent with a number of previous numerical studies, including Battaglia et al. (2012), Pike et al. (2014) and Sembolini et al. (2014).

$Y_{500} - M_{500}$ normalisation: In the middle panel of Fig. 3.9 we plot the normalisation as a function of redshift at the pivot point used in our fit ($M_{500} = 2 \times 10^{14} M_\odot$). The observational constraints are consistent, albeit with large uncertainties, and do not seem to prefer one simulation prediction over another. We note that the Andersson et al. (2011) and Planck Collaboration XX (2014) constraints, which are based on X-ray hydrostatic mass estimates, possess additional uncertainty associated with a possible X-ray mass bias. For example, Planck Collaboration XX (2014) assume a flat prior on the mass bias that allows X-ray hydrostatic masses to be biased low by as much as 30 per cent, which they derive from comparisons with simulations. This corresponds to lowering the normalisation by as much as 0.28 dex compared with no bias, which clearly dominates the uncertainty in the

Planck Collaboration XX (2014) normalisation.

Both the FABLE and MACSIS relations evolve positively in normalisation to at least $z \sim 1.2$. In both cases this reflects the evolution in the gas mass–total mass relation (Fig. 3.3). This can be attributed to AGN feedback having a smaller impact on clusters at higher redshift due to their deeper potential wells at fixed mass, as well as AGN feedback having had less time to affect their host clusters. Significant evolution in the normalisation of the $Y_{500} - M_{500}$ relation could have important consequences for constraining cosmological parameters from SZ cluster counts. For example, Planck Collaboration XX (2014) find that the limiting factor in their cosmological analysis is the modelling of the $Y_{500} - M_{500}$ relation, in particular its normalisation. Fortunately, the change in normalisation predicted by FABLE and MACSIS is small compared with current observational uncertainties, especially considering the uncertainty associated with a potential X-ray mass bias. Yet as ongoing and future SZ surveys beat down the statistical uncertainties and weak lensing analyses provide more, unbiased mass estimates, our results suggest that future cosmological studies may have to account for beyond self-similar evolution in the mass calibration.

$Y_{500} - M_{500}$ intrinsic scatter: The simulation predictions for the intrinsic scatter of the $Y_{500} - M_{500}$ relation (right-hand panel of Fig. 3.9) are consistent with the observational constraints of Andersson et al. (2011), Planck Collaboration XX (2014) and Nagarajan et al. (2018). At $z \sim 0.5$ the simulations and observations agree on a value of ~ 0.09 dex. At $z = 0$ the FABLE and MACSIS values ($0.10^{+0.02}_{-0.01}$ and $0.10^{+0.01}_{-0.01}$ dex respectively) are slightly higher than Planelles et al. (2017) (0.067 dex). The latter is in better agreement with previous simulation works such as Pike et al. (2014) (0.034) and Le Brun et al. (2016) (≈ 0.04) as well as the values of 0.039 and 0.075 ± 0.01 observed by Arnaud et al. (2007) and Planck Collaboration XX (2014) respectively, although FABLE and MACSIS are consistent with the values of $0.08^{+0.06}_{-0.04}$ and 0.07–0.15 observed by Nagarajan et al. (2018) and Sereno et al. (2015b), respectively. Note that we measure a smaller intrinsic scatter in higher mass haloes, as do Planelles et al. (2017). For example, when limiting our sample to masses $M_{500} > 10^{14} M_{\odot}$ at $z = 0$ we find that the intrinsic scatter drops from $0.10^{+0.02}_{-0.01}$ to $0.07^{+0.01}_{-0.03}$.

The simulations make quite different predictions for the redshift evolution of the intrinsic scatter. MACSIS predicts a roughly redshift-independent scatter while FABLE and Planelles et al. (2017) find a mild increase or decrease with increasing redshift, respectively. The difference is most likely related to the sample selection, in particular the mass distribution. Indeed, we find that limiting our sample

to higher mass haloes lowers the intrinsic scatter preferentially at low redshifts, which reduces the apparent redshift evolution. Similarly, Planelles et al. (2017) predict a constant scatter with redshift when focusing on massive clusters.

3.4.3 *Predicted SZ cluster counts*

In this section we investigate how the choice of $Y_{500} - M_{500}$ relation and its redshift evolution affects predictions for the number of clusters detected in an SZ-selected survey. We are motivated by several ongoing and future SZ surveys such as SPT-3G (Benson et al., 2014; Bender et al., 2018), Advanced ACTpol (Henderson et al., 2016) and CMB-S4 (Abazajian et al., 2016), which are expected to vastly expand the number of known clusters and groups. For example, Benson et al. (2014) predict that the ongoing SPT-3G survey will detect ~ 5000 clusters over 2500 deg^2 compared with ~ 500 detected in SPT-SZ, while the proposed CMB-S4 experiment could identify approximately 45,000 to 140,000 clusters depending on the instrument configuration (Abazajian et al., 2016).

Predicting the expected number of clusters for a given survey as a function of redshift requires knowledge of the limiting halo mass, which can then be applied to the redshift-dependent cluster mass function. For an SZ-selected survey the predicted cluster counts therefore depend sensitively on the assumed relationship between the SZ signal and halo mass. This relation can be constrained from current surveys, however the uncertainties are large and results can vary significantly between studies, a few examples of which we have shown in the previous section (for a thorough overview of recent constraints on the $Y_{500} - M_{500}$ relation we refer to section 8.1 in Nagarajan et al. 2018). In addition, the increased sensitivity of future SZ surveys will enable them to detect lower mass clusters than are present in current observational samples. For example, Benson et al. (2014) predict that SPT-3G will detect clusters to a lower mass limit of $M_{500} \sim 10^{14} M_{\odot}$ compared to $\sim 3 \times 10^{14} M_{\odot}$ for SPT-SZ. As a result, the predicted cluster counts are dependent on the extrapolation of the SZ signal–total mass relation to lower masses than existing surveys are able to probe. Extrapolation from higher masses is particularly uncertain because the effects of feedback are expected to be greater in the low mass regime. Furthermore, many of these new clusters will be at higher redshift than existing samples so that the predicted cluster counts depend also on the assumed redshift evolution of the relation. Here we use some of the $Y_{500} - M_{500}$ relations described in the previous section to study how such aspects can affect the predicted number of clusters as a function of redshift.

3.4.3.1 *Calculating cluster counts*

We discuss our procedure for calculating cluster counts for a given $Y_{500} - M_{500}$ relation in detail below. To summarise, we choose a redshift-dependent SZ flux threshold, $Y_{500,\mathrm{lim}}$, obtain the corresponding threshold mass from the $Y_{500} - M_{500}$ relation and calculate the number of clusters above this mass in redshift bins of width $\delta z = 0.1$. We ignore measurement errors for simplicity but do take into account intrinsic scatter about the $Y_{500} - M_{500}$ relation. We compare the FABLE, MACSIS and Planelles et al. (2017) simulations using their respective best-fitting $Y_{500} - M_{500}$ relations as a function of redshift and compare to the observed $Y_{500} - M_{500}$ relations from Nagarajan et al. (2018) and Andersson et al. (2011) assuming self-similar redshift evolution.

In reality, clusters are not selected on the integrated SZ flux, such as Y_{500}, but on a related quantity such as the signal-to-noise ratio. In the case of the SPT-SZ sample (Bleem et al., 2015) and the SPT-3G predictions (Benson et al., 2014), clusters are selected above a threshold value for the SZ detection significance, ζ, which is a measure of the signal-to-noise across all filter scales (Vanderlinde et al., 2010). Unfortunately, a fixed threshold in ζ does not necessarily correspond to an integrated SZ flux threshold, $Y_{500,\mathrm{lim}}$, that is constant with redshift. For example, confusion with primary CMB fluctuations and atmospheric noise suppresses the detection significance of low-redshift clusters (e.g. Vanderlinde et al. 2010) while the detectability of high-redshift or low-mass clusters depends sensitively on the instrument resolution. Andersson et al. (2011) derive a relationship between ζ and Y_{500} using simulated SPT observations and a subsample of SPT-SZ clusters. Their relation implies that, for a fixed ζ threshold, the corresponding $Y_{500,\mathrm{lim}}$ varies with redshift as $E(z)^{-0.25}$. In the following analysis we therefore assume that $Y_{500,\mathrm{lim}}$ scales as $E(z)^{-0.25}$ and use the SPT-SZ 2500 deg^2 results as a baseline for comparison.

We investigate two values for $Y_{500,\mathrm{lim}}$, which we refer to as 'high' and 'low' detection thresholds in the following. In Section 3.4.3.2 we test a high threshold flux of $Y_{500,\mathrm{lim}} = 3.35 \times 10^{-5}\ E(z)^{-0.25}$ Mpc^2. This value was chosen such that the total number of clusters at $z \leq 1.8$ is approximately equal to the total number of detected clusters in the SPT-SZ 2500 deg^2 survey, assuming the $Y_{500} - M_{500}$ relation of Andersson et al. (2011) and our fiducial cosmology. We effectively normalise the cluster counts to the Andersson et al. (2011) relation in this way because, to our knowledge, Andersson et al. (2011) provide the most reliable estimate of the $Y_{500} - M_{500}$ relation underlying the full SPT-SZ cluster sample[2]. In Section 3.4.3.3 we

[2]Saliwanchik et al. (2015) use a novel method to derive Y_{500} for a similar sample of clusters

investigate a low detection threshold of $Y_{500,\mathrm{lim}} = 10^{-5}\ E(z)^{-0.25}$ Mpc^2. Assuming the Andersson et al. (2011) $Y_{500} - M_{500}$ relation this corresponds to ≈ 5000 clusters at $z < 2$ for a 2500 deg^2 survey, equivalent to the total number of predicted clusters for SPT-3G over the same survey area (Benson et al., 2014).

We calculate the halo mass threshold corresponding to our chosen $Y_{500,\mathrm{lim}}$ at the centre of each $\delta z = 0.1$ redshift bin for each $Y_{500} - M_{500}$ relation. For the observational constraints we assume self-similarity out to $z = 1.8$ while for the simulations, whose best-fitting parameters are reported at the redshifts shown in Fig. 3.9, we calculate the mass threshold at each reported redshift and interpolate on these masses where necessary. Note that we take into account the different values of H_0 assumed by different studies but do not otherwise account for their differing cosmologies.

We determine the number of clusters above the evolving mass threshold as a function of redshift as follows. For each redshift bin we calculate a halo mass function with bins of width 0.01 dex at the central redshift. We generate the mass function using the `hmf` code (Murray et al., 2013a) with a Tinker et al. (2008) fitting function and a transfer function computed using the Code for Anisotropies in the Microwave Background (CAMB; Lewis et al. 2000) for which we use the default `hmf` parameters. We take into account intrinsic scatter about the best-fitting relation by assuming that haloes are normally distributed in $\log_{10}(Y_{500})$ at fixed mass with standard deviation equal to the intrinsic scatter. Then, the number density of haloes in each halo mass bin is scaled by the fraction of clusters that are expected to lie above $Y_{500,\mathrm{lim}}$ at the given halo mass. For the simulation relations the intrinsic scatter at a given redshift is linearly interpolated from the reported values[3] while for the observed relations we assume that the intrinsic scatter is independent of redshift. Finally, for each redshift bin we take the integrated number density of haloes above the appropriate halo mass threshold and multiply it by the volume of the redshift bin as seen in a 2500 deg^2 survey.

3.4.3.2 *High SZ detection threshold*

The left-hand panel of Fig. 3.10 shows the cluster counts as a function of redshift corresponding to our high detection threshold ($Y_{500,\mathrm{lim}} = 3.35 \times 10^{-5}\ E(z)^{-0.25}$ Mpc^2). There is a wide spread in the predicted cluster counts as a function of redshift for the different $Y_{500} - M_{500}$ relations. Indeed, the pre-

to Andersson et al. (2011). They do not derive a best-fitting $Y_{500} - M_{500}$ relation, however they compare their results with overlapping data from Andersson et al. (2011) and find consistent results.

[3] In our simulations we find that the intrinsic scatter at the threshold mass is close to the global intrinsic scatter for the mass thresholds under consideration.

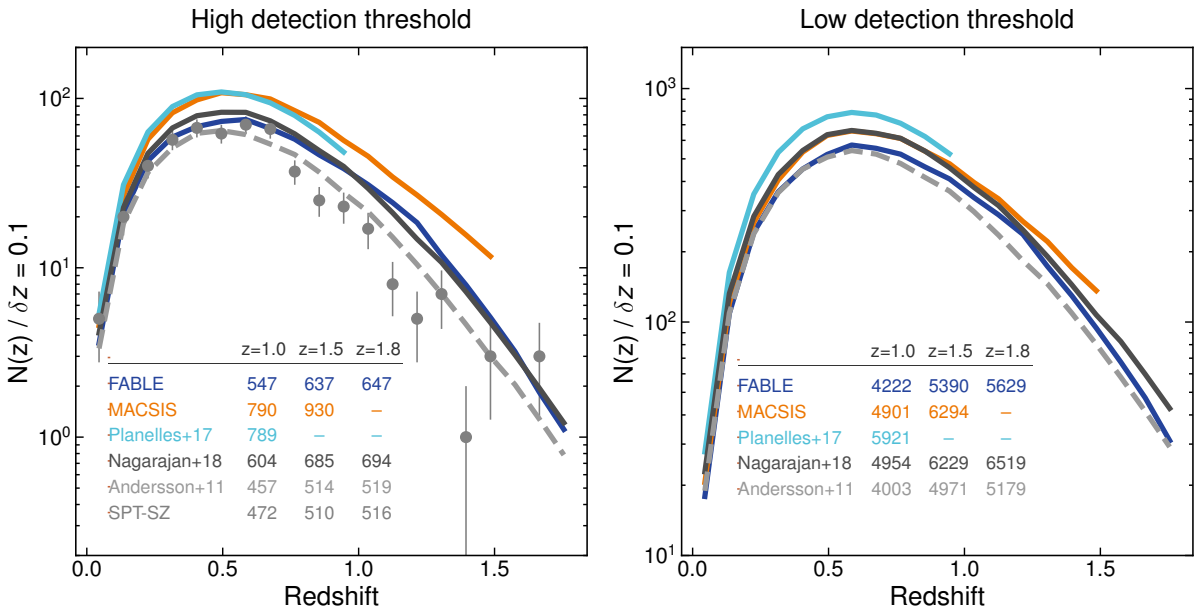


Figure 3.10: The number of SZ-detectable haloes per redshift bin of width $\delta z = 0.1$ over 2500 deg^2 of the sky corresponding to different assumptions for the underlying $Y_{500} - M_{500}$ relation (solid and dashed curves) for two choices of the threshold flux: a high detection threshold, $Y_{500,\mathrm{lim}}\, \mathrm{D}_\mathrm{A}^2 = 3.35 \times 10^{-5}\ E(z)^{-0.25}\ \mathrm{Mpc}^2$ (left), and a low detection threshold, $Y_{500,\mathrm{lim}}\, \mathrm{D}_\mathrm{A}^2 = 10^{-5}\ E(z)^{-0.25}\ \mathrm{Mpc}^2$ (right). Curves correspond to the best-fitting $Y_{500} - M_{500}$ relations of FABLE (dark blue), MACSIS (orange) and Planelles et al. (2017) (light blue), Nagarajan et al. (2018) (solid grey) and Andersson et al. (2011) (dashed grey). Circles with error bars show the cluster counts of the 2500 deg^2 SPT-SZ survey using the updated redshifts from Bocquet et al. (2018) in redshift bins of width $\delta z = 0.1$ (error bars equal the square-root of the number of clusters). The inset table lists the total number of haloes between $z = 0$ and $z = 1.0, 1.5$ and 1.8 for each relation.

dicted number of clusters below $z = 1$ varies from ≈ 600 for Nagarajan et al. (2018) (assuming self-similar evolution) and FABLE to ≈ 800 for the MACSIS and Planelles et al. (2017) simulations. This is largely driven by small differences in the normalisation of the $Y_{500} - M_{500}$ relations (middle panel of Fig. 3.9), plus a small dependence on the slope. The difference in intrinsic scatter between relations is small enough that it has little effect on the cluster counts.

Planelles et al. (2017) predict a relatively high normalisation for the $Y_{500} - M_{500}$ relation, which results in a relatively low mass threshold and a corresponding increase in the number of clusters. In other words, clusters in the Planelles et al. (2017) simulations possess a comparatively high SZ flux at fixed mass, which translates to a higher detection rate in an SZ-selected survey. In the opposite sense, FABLE shows a slightly lower normalisation (~ 0.15 dex) and predicts significantly fewer clusters. The MACSIS relation has a similarly low normalisation to FABLE at $z = 0$ according to Fig. 3.9, however the steepness of their relation means that

the mass threshold corresponding to $Y_{500,\mathrm{lim}} = 3.35 \times 10^{-5}$ Mpc^2 is closer to that of Planelles et al. (2017) than FABLE and hence they predict a similar number of clusters. Similar reasoning applies to the FABLE and Nagarajan et al. (2018) relations, for which the predicted cluster counts agree very closely.

The shape of the cluster counts distribution with redshift is similar between the relations. This reflects the general lack of redshift evolution in the relations beyond that expected in the self-similar model, as evidenced by the flatness of the curves in Fig. 3.9. FABLE and MACSIS predict a mild positive evolution in the normalisation relative to self-similarity, which lowers the mass threshold corresponding to $Y_{500,\mathrm{lim}}$ with increasing redshift thereby slightly boosting the number of clusters at high redshift.

All of the $Y_{500} - M_{500}$ relations predict a larger number of clusters than the Andersson et al. (2011) relation. If the Andersson et al. (2011) relation accurately describes the $Y_{500} - M_{500}$ relation of the SPT-SZ sample then this suggests that cluster counts predicted from SPT-SZ data alone, such as the SPT-3G predictions, may be underestimated. On the other hand, the Andersson et al. (2011) relation may be biased compared with the full SPT-SZ sample since the Andersson et al. (2011) clusters are on average more massive. As such the best-fitting relation may not accurately describe the relationship between Y_{500} and M_{500} at masses close to the lower mass threshold of the SPT-SZ survey, which could have a significant impact on the predicted cluster counts given the steepness of the cluster mass function.

In the left-hand panel of Fig. 3.10 we also plot cluster counts from the 2500 deg^2 SPT-SZ survey binned by redshift (Bocquet et al., 2018). We caution that comparison to these data is dependent on the value of $Y_{500,\mathrm{lim}}$, which we have chosen such that the Andersson et al. (2011) $Y_{500} - M_{500}$ relation yields approximately the same number of clusters as SPT-SZ at $z \leq 1.8$. Therefore, comparing the predictions (curves) to the SPT-SZ counts (error bars) only describes the level of agreement with the Andersson et al. (2011) prediction (dashed line) and not the actual SPT-SZ data, at least in terms of their normalisation.

The shape of the Andersson et al. (2011) curve is a reasonable match to the SPT-SZ counts, however there is a slight underestimate (overestimate) at low (high) redshifts. This is also the case when using the ΛCDM cosmological constraints derived from the SPT-SZ cluster data in Bocquet et al. (2018). There are a number of possible explanations for this discrepancy. Firstly, the prediction implicitly assumes that all haloes with $Y_{500} \geq Y_{500,\mathrm{lim}}$ in a given redshift bin are detected. However, for a real SZ survey the completeness will depend on aspects of the observations such as the noise level and instrument resolution as well as

the intrinsic scatter in the $\zeta - Y_{500}$ relation, which we do not take into account here. Secondly, the discrepancy could be due to our choice of fitting function. Cosmological analysis of the SPT-SZ cluster data in de Haan et al. (2016) and Bocquet et al. (2018) also assume the Tinker et al. (2008) fitting function, however there are a number of other fits in the literature and these can vary significantly depending on, for example, the size and resolution of the simulations or the choice of halo-finder (e.g. Jenkins et al. 2001; Courtin et al. 2010; Crocce et al. 2010; Angulo et al. 2012; Ishiyama et al. 2015). Thirdly, the redshift trend of cluster counts is also fairly sensitive to the redshift evolution of the halo mass threshold. For example, assuming that $Y_{500,\mathrm{lim}}$ is constant with redshift brings the Andersson et al. (2011) curve into better agreement with the SPT-SZ counts compared to the $E(z)^{-0.25}$ scaling, which suggests that the redshift dependence of the $\zeta - Y_{500}$ relation may be even weaker than the Andersson et al. (2011) estimate. On the other hand, the halo mass threshold depends not only on the redshift evolution of $Y_{500,\mathrm{lim}}$ but also the redshift evolution of the normalisation of the $Y_{500} - M_{500}$ relation. In our analysis these choices are degenerate with one another as long as the slope of the $Y_{500} - M_{500}$ relation is independent of redshift, as it is in the self-similar expectation. Hence the discrepancy between the SPT-SZ counts and the Andersson et al. (2011)–based prediction could also be resolved if the normalisation of the $Y_{500} - M_{500}$ relation evolves less rapidly with redshift than the self-similar expectation of $E(z)^{-2/3}$. This is not supported by the simulation predictions however, which yield either a redshift-independent normalisation consistent with self-similarity as in Planelles et al. (2017) or, in the case of FABLE and MACSIS, a normalisation that evolves with redshift faster than self-similar.

3.4.3.3 *Low SZ detection threshold*

In this section we test a lower flux threshold of $Y_{500,\mathrm{lim}} = E(z)^{-0.25}\, 10^{-5}$ Mpc^2 (right-hand panel of Fig. 3.10), which corresponds to a much larger number of clusters (~ 5000) comparable to the expected return of SZ surveys such as SPT-3G.

The predicted cluster counts are more sensitive to differences in the slope of the relations than for the high detection threshold because $Y_{500,\mathrm{lim}}$ corresponds to a mass much further from the pivot point. In fact, the Andersson et al. (2011) and Nagarajan et al. (2018) clusters all lie above this threshold and so the predicted cluster counts rely on the extrapolation of the fit to lower masses. For example, MACSIS predicts a significantly steeper slope than the other $Y_{500} - M_{500}$ relations so that the mass corresponding to this lower detection threshold is relatively high. Therefore, MACSIS predicts ~ 1000 fewer clusters than Planelles et al. (2017) for the low SZ detection threshold whereas their predictions were similar for the

high detection threshold. Similarly the Nagarajan et al. (2018) relation yields significantly more clusters than the FABLE relation due to their slightly shallower slope. This demonstrates how relatively small changes in the slope of the best-fitting $Y_{500} - M_{500}$ relation can have a significant impact on the expected cluster counts, especially as we push towards lower masses with future surveys.

The mild positive evolution in the normalisation of the FABLE and MACSIS relations has a small but appreciable effect on the shape of the cluster counts distribution. For example, the cluster counts predicted by the FABLE relation are similar to that of the Andersson et al. (2011) relation at $z \lesssim 0.5$, however the increase in normalisation with increasing redshift in the FABLE relation yields relatively more clusters at higher redshift. Similar reasoning applies to the cluster counts predicted by the MACSIS relation, which yields fewer clusters than the Nagarajan et al. (2018) relation at $z < 0.5$ where their normalisation is lower and more clusters at $z > 0.5$ where they predict a higher normalisation.

The total number of clusters varies considerably between the relations, ranging from ~ 4000 for FABLE and Andersson et al. (2011) to ~ 6000 for Planelles et al. (2017) at $z \leq 1.0$. Similarly large differences are also found at higher redshift. This highlights the need for improvements in the measurement of the $Y_{500} - M_{500}$ relation, not only for predicting the outcomes of future SZ surveys but also for taking advantage of such surveys for probing cosmology. Indeed, cosmological constraints from current surveys are primarily limited by uncertainties in the calibration of the SZ signal–total mass relation (e.g. Sehgal et al. 2011; von der Linden et al. 2014b; Bocquet et al. 2015; Planck Collaboration XXIV 2016). Furthermore, simulations such as FABLE and MACSIS hint towards a mild redshift evolution in the relations beyond the self-similar expectation that may prove important to the next generation of SZ experiments, such as CMB-S4, which will find clusters out to the highest redshifts where they exist.

3.5 CONCLUSIONS

We have studied the redshift evolution of the X-ray and SZ scaling relations in the FABLE suite of cosmological hydrodynamical simulations. The simulations are performed using the AREPO moving-mesh code with a set of physical models based on those of Illustris (Genel et al., 2014; Vogelsberger et al., 2014a; Sijacki et al., 2015) but with improved modelling of supernovae and AGN feedback (see Chapter 2 for details). For this study we have greatly expanded our sample of simulated galaxy groups and clusters by performing an additional 21 zoom-in simulations in addition to the original 6 used in Chapter 2. Our extended suite

of 27 high-resolution zoom-in simulations spans a wide halo mass range, from low-mass groups ($\sim 3 \times 10^{13} M_{\odot}$) to massive clusters ($\sim 3 \times 10^{15} M_{\odot}$).

Using our expanded sample we have investigated six scaling relations: gas mass–total mass, total mass–temperature, $Y_{\rm X}$–total mass, X-ray luminosity–total mass, X-ray luminosity–temperature and the SZ flux–total mass relation. First we examined the reliability of our model by comparing the predicted scaling relations to observations at intermediate to high redshift ($z \lesssim 1$; Sections 3.3.1 and 3.4.1), in extension to the $z = 0$ comparison presented in Chapter 2. Subsequently we investigated the redshift evolution of the slope, normalisation and intrinsic scatter of the scaling relations out to $z \approx 2$ in comparison to other recent simulation predictions (Sections 3.3.2 and 3.4.2.2). We find significant deviations from the simple self-similar expectation for all of the relations examined here (Fig. 3.3-3.5 and Fig. 3.9). These predictions relate directly to the outcomes of future experiments such as CMB-S4 (Abazajian et al., 2016) and the *Athena* X-ray observatory (Nandra et al., 2013). In particular, our mock X-ray analysis mimics observations with the planned *Athena* X-IFU instrument, which we expect to find significant deviations from self-similarity in the evolution of the X-ray scaling relations out to $z \sim 2$. This could have important implications for cluster cosmology, particularly as ongoing and future cluster surveys push to higher redshift. We illustrate this point in Section 3.4.3 by comparing predicted cluster counts from an SZ-selected survey for different assumptions about the underlying SZ flux–total mass relation. We provide a summary of our main results below.

- At $z = 0.4$ and $z = 1$, the gas mass–total mass, $Y_{\rm X}$–total mass and X-ray luminosity–total mass relations are in good agreement with observational constraints based on X-ray hydrostatic mass estimates (Fig. 3.1 and 3.2). On the other hand, the X-ray luminosity–spectroscopic temperature relation lies on the upper end of the observed scatter. Comparison to observations based on weak lensing masses as opposed to X-ray hydrostatic masses suggests that FABLE clusters possess slightly overestimated gas masses and X-ray luminosities at fixed total mass, although we find no evidence for a significant redshift dependence in the offset.

- We find that the slopes of the relations are in good agreement with the majority of observations at low redshift. Furthermore, the slopes deviate significantly from the self-similar predictions in all cases and at all redshifts ($0 < z < 1.8$; left-hand panels of Fig. 3.3-3.5 and Fig. 3.9). The same qualitative result is found in the recent studies of B17, for the MACSIS simulations, and T18, for a set of simulations with AGN feedback. The predicted slopes

are consistent with the MACSIS and T18 predictions at $z \lesssim 2$, with the exception of the gas mass–total mass and $Y_{\rm X}$–total mass relations for which T18 predict somewhat shallower slopes.

- The scaling relations of gas mass, $Y_{\rm X}$, SZ flux and X-ray luminosity with total mass are all steeper than the self-similar expectation. This is largely due to the effects of non-gravitational physics, such as star formation and feedback. In particular, AGN feedback can expel gas with greater efficiency from lower mass haloes due to their shallower potentials wells. The fact that T18 predict a shallower gas mass–total mass slope than FABLE and MACSIS (Fig. 3.3) suggests that gas expulsion via AGN feedback is less efficient in their simulations, at least at $z \lesssim 2$. This may stem from differences in the frequency of thermal energy injection by AGN feedback, which is continuous in the T18 model but operates on a duty cycle in FABLE and MACSIS.

- The total mass–temperature slope based on the spectroscopic temperature is only marginally steeper than self-similar due to spectroscopic temperature biases at the low- and high-mass end of the relation. On the other hand, the mass-weighted temperature relation is significantly steeper than self-similar, in good agreement with MACSIS and T18 (Fig. 3.4). This is due to the combined action of radiative cooling, star formation and AGN feedback, which raises the temperature of the ICM and acts with greater efficiency in lower mass haloes.

- We find no strong evidence for a redshift evolution in the slopes of the relations, with the exception of the X-ray luminosity-based relations for which we see a mild steepening with increasing redshift (Fig. 3.5). We also find a mild increase in slope with decreasing redshift at $z \lesssim 0.6$ for the gas mass, $Y_{\rm X}$ and X-ray luminosity-based relations, which may be driven by the increasing prevalence of radio-mode AGN feedback in our simulations.

- The normalisation of the scaling relations evolves positively with respect to self-similarity in all cases (middle panels of Fig. 3.3-3.5 and Fig. 3.9). The positive evolution of the gas mass, X-ray luminosity and SZ flux-based relations is because haloes of a given mass are denser at higher redshift, which raises the energy required to expel gas beyond r_{500}. The evolution of the total mass–temperature relation is due to increasing non-thermal pressure support from kinetic motions in the ICM with increasing redshift, as found in Le Brun et al. (2016) for the cosmo-OWLS simulations. Contrary to the results of T18, this is not sufficient to offset the evolution in the gas

mass at fixed mass, and so the Y_X–total mass relation evolves positively with respect to self-similarity. The spectroscopic temperature bias increases with increasing redshift due to redshifting of the low-energy X-ray emission below the X-ray bandpass. This largely offsets the otherwise positive evolution of the total mass–temperature and X-ray luminosity–temperature relations so that they appear to evolve self-similarly when based on spectroscopic temperatures as opposed to mass-weighted temperatures.

- The intrinsic scatter is in good agreement with the majority of observational and theoretical constraints at low redshift, although in some cases it is biased high by the scattering of low-mass objects in our comparatively low-mass sample. FABLE and MACSIS predict a larger scatter in the gas mass at fixed total mass compared with T18, which may be related to differences in the duty cycle of AGN feedback. The scatter tends to decrease with increasing redshift (right-hand panels of Fig. 3.3-3.5 and Fig. 3.9), consistent with Le Brun et al. (2016) for a large, volume-limited sample. This is likely a combination of the decreasing efficiency of AGN feedback with redshift and the increased influence of radio-mode AGN feedback at $z \lesssim 1$ in our model.

- The scaling relation between SZ flux and total mass is in good agreement with *Planck* and SPT-SZ clusters out to $z \sim 0.8$ and $z \sim 1.2$, respectively (Fig. 3.6 and 3.7). The $Y_{500} - M_{500}$ relation at $z = 0$ also agrees well with local cluster observations (Fig. 3.8). At group scales, the simulation predictions favour the relation of Wang et al. (2016) for locally brightest galaxies (recalibrated from Planck Collaboration XI 2013) over that of Lim et al. (2018) for local galaxy groups.

- We show that the number and redshift distribution of clusters expected from an SZ-selected survey depends sensitively on the assumed slope, normalisation and redshift evolution of the $Y_{500} - M_{500}$ relation (Fig. 3.10). Relatively small differences in the normalisation (~ 0.1 dex) can result in predicted cluster counts that differ by more than a thousand for an SZ-selected survey similar to SPT-3G (with an expected total of ~ 5000 clusters). Furthermore, the steeper than self-similar $Y_{500} - M_{500}$ relations of FABLE and MACSIS yield relatively fewer clusters at low redshift compared with the close to self-similar prediction of Planelles et al. (2017), whereas the mild positive evolution in their normalisations leads to a relatively higher number of high-redshift clusters. This has important consequences for SZ cluster cosmology,

which depends on our ability to link observed SZ counts to the underlying cluster mass function.

As ongoing and future cluster surveys detect lower mass systems at higher redshifts, our ability to leverage this new data to probe cosmology will depend on our understanding of the mass and redshift dependence of the observable–mass relations. With FABLE we predict significant deviations from the self-similar expectation in terms of the slope of the relations and the redshift evolution of their normalisation. Fortunately for observational studies, the assumption of a redshift-independent slope seems to be robust for the relations examined here. On the other hand, none of the relations we have examined evolves self-similarly in normalisation, including commonly-used mass proxies such as the $Y_{\rm X}$ parameter. This could have a significant impact on the expected number of high-redshift clusters from future surveys (see e.g. Section 3.4.3) and, relatedly, our ability to constrain cosmology using cluster abundances. Furthermore, the intrinsic scatter of the scaling relations tends to decrease with increasing redshift and increasing mass, which implies the need for a more complex parametrization of the intrinsic scatter than the mass- and redshift-independent scatter that is assumed in most current observational studies.

4 TOTAL BARYON CONTENT AND BCG PROPERTIES

In this chapter I study the total baryon content of galaxy groups and clusters in the FABLE simulations, as well as the mass, size and shape of their central brightest cluster galaxies (BCGs). The total gas mass and total stellar mass are in good agreement with observations as a function of halo mass, although comparison to observational constraints based on weak lensing mass estimates as opposed to X-ray hydrostatic masses implies that FABLE groups and clusters are too gas-rich. The gas and stellar mass of intermediate mass clusters ($M_{500} \gtrsim 3 \times 10^{14} M_{\odot}$) show negligible redshift evolution at $z \lesssim 1$, in agreement with recent findings from Sunyaev-Zel'dovich (SZ)-selected cluster samples. BCGs in FABLE have somewhat higher stellar mass at fixed cluster mass than observed (by ~ 0.2–0.3 dex), similar to the recent IllustrisTNG and C-EAGLE simulations. The redshift evolution of the stellar mass in simulated BCGs is intermediate between observational studies that measure a mild (~ 35 per cent) or rapid (~ 100 per cent) BCG mass growth between $z = 1$ and $z = 0$. The stellar mass profiles of FABLE BCGs and the surrounding intracluster light (ICL) are similar to observed profiles at $z \approx 0$. Beyond the typical surface brightness limit of the observations ($\gtrsim 100$ kpc) the simulated profiles are highly extended, following a close to power-law shape out to several hundred kpc and with a slightly shallower slope at $z = 0.2$ compared with $z = 1$. A Sérsic model fit to the stacked profiles underestimates the stellar mass density at $\gtrsim 100$ kpc and biases the inferred size growth with redshift. Multi-component models such as the triple Sérsic profile are better able to capture these outer regions, provided the outer radius of the fit is sufficiently large.

The work presented in this chapter is currently being prepared for publication in MNRAS. The analyses and discussion are entirely my own work.

4.1 GLOBAL BARYONIC PROPERTIES

Situated at the nodes of the cosmic web, galaxy cluster formation is characterised by continual accretion of dark matter, gas and stars along connecting filaments. This process is largely dominated by gravity, which couples equally to all matter types. As such, we expect galaxy clusters to be fair samples of the matter content of the Universe at large. Yet there exist a number of key physical processes that may act to suppress total baryon fractions compared to the universal value, for example, feedback from AGN and massive stars. These are of particular importance in less massive systems, such as galaxy groups, where non-gravitational processes begin to prevail over gravity due to the shallower potential wells of these objects. In addition, the partitioning of the baryons into gas in the hot ICM and stars provides information about cluster assembly and the efficiency of star formation integrated over the cluster's formation history.

Indeed, studies that have measured the baryon fractions as a function of total mass have shown that feedback has a greater impact on low-mass than high-mass haloes and that star formation efficiency is higher in lower mass systems (e.g. Lin et al. 2003; Gonzalez et al. 2007; Giodini et al. 2009; Gonzalez et al. 2013). In this vein, in Sections 4.1.1 and 4.1.2 we study how the total mass of gas and stars in FABLE groups and clusters depends on halo mass with comparison to a range of observational data, as well as predictions from other recent simulations. In addition, with the recent rise in the number of high-redshift clusters detected via the Sunyaev-Zel'dovich (SZ) effect, some studies have begun to tackle the question of how the baryonic components of clusters evolve with redshift (e.g. Chiu et al. 2016a,b). Recently, Chiu et al. (2018) analysed an SZ-selected sample of clusters at $0.2 < z < 1.25$ with X-ray, optical and near-infrared follow-up data to show that the baryon content of clusters has remained almost unchanged over the past 9 Gyr of cosmic evolution. In Sections 4.1.3.1 and 4.1.3.2 we test this finding by directly comparing the redshift evolution of the total gas and stellar content of FABLE groups and clusters to the results from Chiu et al. (2018).

4.1.1 *ICM mass to total mass relation*

In the left-hand panel of Fig. 4.1 we show the total gas mass within r_{500} as a function of halo mass for FABLE groups and clusters (blue diamonds) in comparison to data (grey symbols). The data from Eckert et al. (2016) are based on weak lensing mass estimates while Maughan et al. (2008), Sun et al. (2009), Gonzalez et al. (2013), Sanderson et al. (2013) and Lovisari et al. (2015) use X-ray

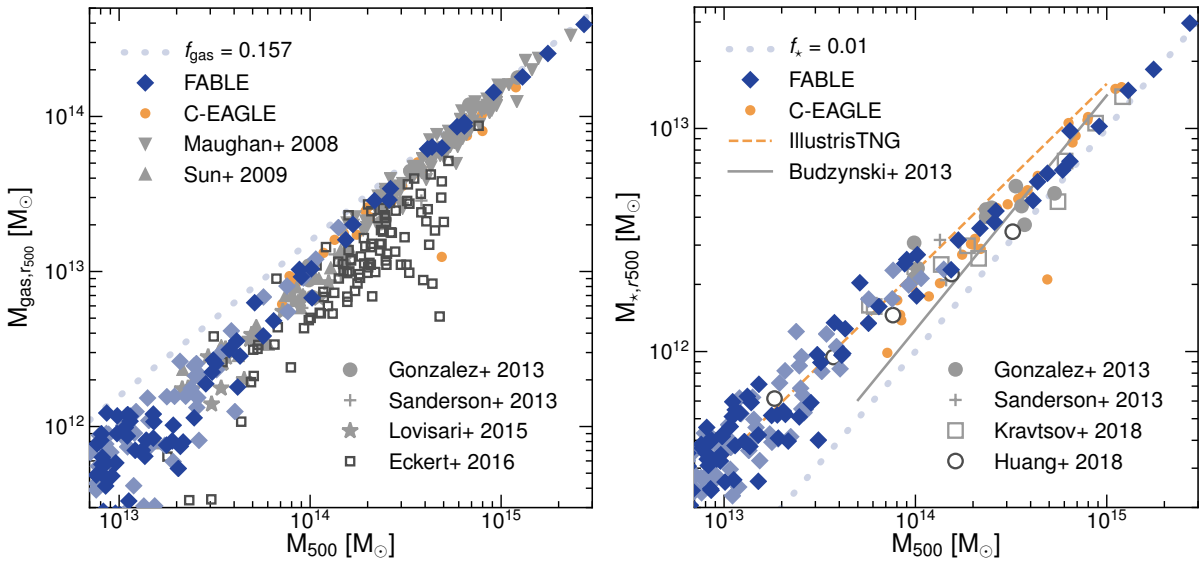


Figure 4.1: The total mass of gas (left) and stars (right) within r_{500} at $z = 0$ as a function of total mass for the FABLE systems (diamonds) compared to observational data (grey symbols) as well as C-EAGLE clusters (orange circles) and the best-fitting relation from IllustrisTNG (dashed line). All of the simulation results use total masses measured directly from the simulation. Dark blue diamonds correspond to haloes within our cosmological box and the main halo of each zoom-in simulation. Light blue diamonds indicate secondary haloes in the high-resolution region of the zoom-in simulations. Grey symbols show observed groups and clusters, which are dark or light grey for data based on weak lensing mass estimates or X-ray hydrostatic masses, respectively. In the right-hand panel, grey open circles indicate the median relation of Huang et al. (2018) and the grey solid line shows the best-fitting relation from the stacked analyses of Budzynski et al. (2013). Thick dotted lines correspond to a constant gas or stellar mass fraction of 15.7 per cent and 1 per cent, respectively. The former is approximately the cosmic baryon fraction in our fiducial cosmology.

hydrostatic masses. We also compare to the recent C-EAGLE cluster simulations (orange circles; Barnes et al. 2017b; Bahé et al. 2017), a suite of hydrodynamical zoom-in simulations of 30 galaxy clusters performed with the same resolution and galaxy formation physics as the EAGLE simulation (Schaye et al., 2015; Crain et al., 2015). For C-EAGLE and FABLE we use the halo masses measured directly from the simulation. The gas masses of C-EAGLE clusters are estimated from the mock X-ray pipeline described in Barnes et al. (2017b). For FABLE haloes we calculate the total gas mass within a radius r_{500} of the gravitational potential minimum of the halo. This definition includes cold as well as hot, X-ray emitting gas, however in Fig. 2.5 we showed that cold gas makes a negligible contribution to the total gas mass on group and cluster mass scales ($M_{500} \gtrsim 10^{13} M_\odot$).

The slope of the simulated relation is in good agreement with the majority of the observational constraints. All of the relations diverge significantly from the

self-similar expectation in the absence of non-gravitational physics, which predicts a constant gas mass fraction at all mass scales. Indeed, on the scale of low-mass clusters and groups ($M_{500} \lesssim 3 \times 10^{14} M_{\odot}$) the ratio of the total gas mass to the total mass of the cluster is significantly lower than the cosmic baryon fraction ($f_b \equiv \Omega_b/\Omega_m \approx 0.157$; dotted line in the left-hand panel of Fig. 4.1), whereas at the high mass end of the relation the slope is consistent with a constant gas mass fraction approaching the cosmic value. This presents strong evidence for the impact of non-gravitational processes, such as AGN feedback, on the total gas content of all but the most massive clusters.

In terms of normalisation, the simulated relation is in very good agreement with the observational data based on X-ray hydrostatic mass estimates. This is partly by design, since the strength of AGN feedback in the FABLE model was calibrated to reproduce the gas mass fractions of galaxy groups with $M_{500} \lesssim 10^{14} M_{\odot}$ (see Chapter 2). Even so, it is worth pointing out that the match to observations at cluster-scale masses ($M_{500} > 10^{14} M_{\odot}$) was not guaranteed by the calibration, as the deeper gravitational potential wells of these systems increases the energy required for AGN feedback to eject gas beyond r_{500}.

Conversely, the FABLE groups and clusters have significantly higher gas mass at fixed total mass compared to data from Eckert et al. (2016), who estimate total mass via a weak lensing-calibrated relation between total mass and X-ray temperature. A similar predicament holds for the C-EAGLE clusters (orange circles in Fig. 4.2), which have similar gas masses to the FABLE systems at fixed halo mass. The difference between the results of Eckert et al. (2016) and those based on X-ray hydrostatic masses can be explained if X-ray masses are biased low compared to weak lensing masses. Indeed, Barnes et al. (2017b) derive X-ray hydrostatic mass estimates for the C-EAGLE clusters and find that they are biased low compared to the true halo mass by ≈ 20 per cent on average. The issue of a X-ray hydrostatic mass bias and its implications for our results have been discussed in detail in Chapter 2. To summarise, we compared the FABLE predictions for a number of X-ray scaling relations at $z = 0$ to data and found that many of the discrepancies of the model relative to observations are consistent with a significant X-ray hydrostatic mass bias in observed samples. In particular, FABLE groups and clusters are too luminous in X-rays at fixed temperature compared to observations. This could stem from FABLE systems being too gas-rich at fixed total mass, which is the primary consequence of a significant X-ray mass bias given that the FABLE model was calibrated to observed gas mass fractions based on X-ray hydrostatic mass estimates.

On the other hand, a large X-ray hydrostatic mass bias ($\gtrsim 30$ per cent) implies

low baryon fractions in clusters that are difficult to reconcile with our current understanding of galaxy cluster formation. For example, the weak lensing-calibrated masses used in Eckert et al. (2016) correspond to a roughly mass-independent hydrostatic mass bias of 28 per cent and a baryon fraction of 0.067 ± 0.008 in $10^{14} M_{\odot}$ haloes within r_{500}. This falls short of the universal baryon fraction in a Planck Collaboration XIII (2016) cosmology ($\Omega_b/\Omega_m \simeq 0.157$) by more than a factor of two. The corresponding baryon depletion factor is $\mathcal{D} \approx 0.57$, where $\mathcal{D} = 1 - f_{\rm bar}/(\Omega_{\rm b}/\Omega_{\rm m})$ and $f_{\rm bar}$ is the cluster baryon fraction. In contrast, simulations predict a much smaller depletion factor, on the order of $\mathcal{D} \sim 0.3$ at $10^{14} M_{\odot}$ (e.g. Planelles et al. 2013; Le Brun et al. 2014; McCarthy et al. 2017). Indeed, the baryon depletion factor for FABLE haloes at $M_{500} \sim 10^{14} M_{\odot}$ is only $\mathcal{D} \sim 0.2$, although this may be underestimated if clusters are too gas-rich. Models with stronger AGN feedback can yield large depletion factors in agreement with the Eckert et al. (2016) results, however these models struggle to reproduce other cluster observables, such as the thermodynamic profiles of the ICM (see e.g. Le Brun et al. 2014). A solution to this problem may be a more sophisticated modelling of AGN feedback in simulations and/or the inclusion of previously neglected physical processes that, perhaps in combination, are able to efficiently lower baryon fractions in groups and clusters from the universal average without overheating or evacuating gas in the core regions. Nevertheless, to reliably calibrate these models to observations we first require a thorough understanding of the mass bias in order to fully constrain the baryon content of clusters as a function of their mass. Looking beyond the calibration, we can then utilise simulations to investigate the degree to which different non-gravitational processes are responsible for the depletion of baryons in clusters.

4.1.2 *Stellar mass to total mass relation*

One of the most commonly-used observational tracers of galaxy clusters is the galaxies themselves. Originally identified as overdensities of galaxies on the sky, clusters typically consist of numerous satellite galaxies surrounding a central brightest cluster galaxy (BCG). In addition, observations have revealed a significant proportion of stars forming a diffuse component of intracluster light (ICL; e.g. Zhang et al. 2018). The total stellar content of a cluster is the sum of these three components (satellites, BCG and ICL) and holds a great deal of information about the galactic accretion and star formation history of the cluster.

In the right-hand panel of Fig. 4.1 we show the total stellar mass of FABLE groups and clusters within r_{500} as a function of total mass in comparison to

observational and theoretical constraints at $z \approx 0$. We limit the comparison to studies that take into account a contribution from the ICL, as this represents a significant fraction of the total stellar mass in galaxy clusters and groups, both in our simulations and in observations (e.g. ≈ 20–50 per cent; Gonzalez et al. 2007; Seigar et al. 2007; Zibetti 2007; McGee & Balogh 2010; Gonzalez et al. 2013, but see also Krick & Bernstein 2007; Burke et al. 2015; Jiménez-Teja et al. 2018 who find smaller ICL fractions of ≈ 5–25 per cent). We compare to data from Budzynski et al. (2013), Gonzalez et al. (2013), Sanderson et al. (2013) and Kravtsov et al. (2018), which use X-ray hydrostatic masses, and the median relation from Huang et al. (2018), which is based on weak lensing mass estimates.

In general the FABLE clusters and groups seem to have formed a realistic total mass of stars at $z = 0$. In massive clusters, approximately 1 per cent of the total mass is in the form of stars (dotted line in Fig. 4.1), whereas lower mass haloes diverge towards higher stellar mass fractions, in agreement with the observations. At the low-mass end ($M_{500} \lesssim 10^{14} M_{\odot}$) we have good agreement with the median relation of Huang et al. (2018) who measure stellar masses for a large sample of galaxies from the Hyper Suprime-Cam (HSC) survey (Aihara et al., 2018) with halo masses measured from weak lensing. Similarly, at the high mass end we have excellent agreement with the clusters studied in Gonzalez et al. (2013), Sanderson et al. (2013) and Kravtsov et al. (2018) with X-ray hydrostatic mass estimates. In their highest mass bin ($M_{500} = 3.2 \times 10^{14} M_{\odot}$), Huang et al. (2018) measure a slightly lower stellar mass than Gonzalez et al. (2013) for clusters of similar halo mass. This may be due to X-ray hydrostatic mass bias, although at a somewhat lower level than that implied by the weak lensing calibrated gas mass fractions of Eckert et al. (2016) discussed in the previous section. Alternatively the offset may be related to the mass-to-light ratios used in these studies to convert the measured luminosities into stellar mass. For example, for a sample of high-redshift ($z \sim 1$) clusters, van der Burg et al. (2014) show that assuming a fixed mass-to-light ratio for all galaxies (as in Gonzalez et al. 2013), rather than deriving it for each galaxy individually based on SED modelling (as in Huang et al. 2018), overestimates the total stellar mass in their clusters by at least a factor of two. Whether this can explain the differences seen at $z \approx 0$ remains unclear however, as the bias is expected to be smaller at lower redshift (van der Burg et al., 2014). If the Gonzalez et al. (2013) stellar masses are indeed biased high at fixed total mass – either due to X-ray mass bias or overestimated mass-to-light ratios – then this would imply that the total stellar masses of FABLE clusters are slightly overestimated.

The same reasoning also applies to C-EAGLE, which predicts similar stellar masses to FABLE. The C-EAGLE galaxy formation model predicts significantly

fewer high-mass galaxies ($M_\star \gtrsim 10^{11} M_\odot$) than FABLE in the field environment at $z \approx 0$ (see the galaxy stellar mass function comparison with EAGLE in Fig. 2.2). This suggests that cluster-specific processes responsible for the suppression of star formation may be more effective in FABLE than in C-EAGLE. We also compare to the best-fitting relation from the IllustrisTNG suite of cosmological magneto-hydrodynamical simulations of galaxy formation, which is based on halo masses measured directly from the simulation (orange dashed line in Fig. 4.1; Pillepich et al. 2018b). The prediction from IllustrisTNG is in good agreement with FABLE and Huang et al. (2018) at the low mass end but lies slightly above the other relations at cluster mass scales ($\gtrsim 10^{14} M_\odot$). This may reflect differences in the field galaxy stellar mass function, for which IllustrisTNG predicts a slightly higher abundance of massive galaxies compared to FABLE and significantly more than EAGLE (Pillepich et al., 2018b).

The simulations predict a fairly tight relation between stellar mass and total mass, with a level of intrinsic scatter comparable to, but slightly lower than, the observational constraints. In particular, a power-law fit to FABLE clusters with $M_{500} > 10^{14} M_\odot$ yields a log-normal intrinsic scatter of $0.06^{+0.02}_{-0.01}$ dex in stellar mass at fixed total mass.[1] This is consistent with the scatter predicted by IllustrisTNG (0.07 dex; Pillepich et al. 2018b) but somewhat smaller than the intrinsic scatter of 0.09 ± 0.05 dex for the Kravtsov et al. (2018) sample and 0.11 ± 0.03 dex for an extended sample including clusters from Gonzalez et al. (2013).

The slope of the FABLE relation varies somewhat on different halo mass scales, being notably shallower in the low-mass cluster regime ($M_{500} \lesssim 5 \times 10^{14} M_\odot$) than for higher mass clusters, where it is consistent with a constant stellar mass fraction. This is in contrast to the C-EAGLE and IllustrisTNG relations, which have a roughly constant slope, but is in good agreement with the Kravtsov et al. (2018) sample. For the mass range in which they overlap, the slope of the FABLE relation is also consistent with Gonzalez et al. (2013) but is somewhat steeper than that of Huang et al. (2018) and shallower than the relation of Budzynski et al. (2013). The latter is significantly steeper than the other observational constraints and implies that group-scale haloes ($M_{500} \lesssim 10^{14} M_\odot$) have a much lower mass of stars than is seen in other studies or predicted by the simulations. The origin of this difference is unclear, although it may be related to the fact that Budzynski et al. (2013) find little to no contribution from the ICL on galaxy group scales. In fact,

[1] Unless specified otherwise, best fits are performed in log-space using the orthogonal BCES method (Akritas & Bershady, 1996) and quoted uncertainties on the best-fitting parameters correspond to the 68 per cent confidence interval estimated from bootstrapping with 10^4 resamples. The intrinsic scatter is computed according to equation 3.2.

they measure an ICL mass fraction that mildly increases with total mass, opposite to Gonzalez et al. (2013) and Kravtsov et al. (2018) who find a decreasing ICL contribution with increasing total mass. The determination of the total stellar mass in the ICL is notoriously difficult due to its diffuse nature, which requires very deep observations and careful consideration of contamination by foreground and background galaxies. In addition, the measured ICL fraction depends sensitively on the method used to separate the light of the BCG from the diffuse ICL, which blend together smoothly in the outer regions of the BCG. The latter point also introduces significant uncertainty when it comes to determining the ICL fraction in simulations, which has been a recurrent theme in previous works (e.g. Conroy et al. 2007; Dolag et al. 2010; Puchwein et al. 2010; Contini et al. 2014; Cooper et al. 2015). We do not attempt to measure the ICL fraction in our simulations for this reason, although we note that the extended stellar mass profiles presented in Section 4.2.4 suggest that there is a significant component of ICL in FABLE clusters out to several hundred kpc.

4.1.3 *Baryonic mass versus redshift*

In Sections 4.1.3.1 and 4.1.3.2 we compare the gas and stellar mass content of FABLE clusters as a function of redshift to results from Chiu et al. (2018), who study an SZ-selected sample of 91 galaxy clusters at $0.2 < z < 1.25$ detected in the 2500 deg^2 SPT-SZ survey. Gas masses are derived from *Chandra* X-ray data and stellar masses from optical photometry obtained in the Dark Energy Survey (Dark Energy Survey Collaboration, 2005, 2016) and near-infrared photometry from the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010) and *Spitzer* (Ashby et al., 2013). Total masses are estimated from the SZ signal using the best-fitting scaling relation for SPT-SZ clusters derived in de Haan et al. (2016). This relation is determined by fitting a ΛCDM cosmology model – with external priors from Big Bang nucleosynthesis calculations (Cooke et al., 2014) and direct measurement of the Hubble parameter (Riess et al., 2011) – to the SPT cluster data set, which includes a subset of clusters with mass estimates derived from a weak lensing-calibrated relation between $Y_{\rm X}$ and total mass (Bocquet et al., 2015; de Haan et al., 2016).

The Chiu et al. (2018) sample, by virtue of its selection on the SZ signal, is approximately mass-limited, with a minimum mass of $M_{500} \approx 3 \times 10^{14} M_{\odot}$ across the full redshift range. For our comparison sample we therefore include all FABLE clusters with a total mass greater than $M_{500} = 3 \times 10^{14} M_{\odot}$ at each redshift. We have verified that this yields a similar range of masses as the observed sample at

each redshift.

Chiu et al. (2018) derive a best-fitting scaling relation linking the gas mass or stellar mass to the total mass and redshift with (redshift-independent) log-normal intrinsic scatter at fixed total mass. All masses are measured within r_{500}. Following Chiu et al. (2018), in Fig. 4.2 we plot the gas mass (left-hand panel) and stellar mass (right-hand panel) as a function of redshift after removing the mass dependence of these quantities using the best-fitting scaling relations. As such, Fig. 4.2 highlights the redshift trend of the gas mass and stellar mass at the pivot point of $M_{500} = 4.8 \times 10^{14} M_{\odot}$. We scale the FABLE quantities – and the data from Gonzalez et al. (2013) with which we complement the low-redshift comparison – by their respective best-fitting mass slopes as indicated in the figure legends. For FABLE we use the best-fitting slopes at $z = 0$ for the mass range in question ($1.02^{+0.02}_{-0.03}$ and $0.92^{+0.03}_{-0.03}$ for the gas mass and stellar mass, respectively), which are consistent with the slopes at higher redshift ($z \leq 1$) to better than 1-sigma. We compare to clusters with $M_{500} > 3 \times 10^{14} M_{\odot}$ from Gonzalez et al. (2013) but use the best-fitting slopes of their full sample as there are too few clusters in this mass range with which to derive the slope. The stellar mass to halo mass slope of Gonzalez et al. (2013) (0.52 ± 0.04) is somewhat shallower than Chiu et al. (2018) (0.80 ± 0.12) and FABLE ($0.92^{+0.03}_{-0.03}$), however this has a negligible effect on the comparison given that their sample spans a small halo mass range, with a median mass close to the pivot point. We note that choosing the best-fitting mass slope of Chiu et al. (2018) for all three datasets does not change our conclusions.

4.1.3.1 *ICM mass redshift trend*

The dashed line in the left-hand panel of Fig. 4.2 shows the best-fitting relation from Chiu et al. (2018), which implies that the gas mass at fixed total mass varies with redshift as $(1+z)^{-0.15 \pm 0.14}$. The redshift trend is therefore mildly negative but statistically consistent with zero. Our results appear to support this finding for the mass range in question. Indeed, the median gas mass of the simulated sample remains roughly constant with redshift up to $z \sim 1$, albeit with a relatively small sample size. This implies that the gas mass at fixed total mass is kept constant via a balance between (1) accretion of gas from the cluster outskirts (beyond r_{500} in this case), which has a gas mass fraction similar to or larger than that of the cluster, (2) accumulation of lower mass objects with smaller gas mass fractions, and (3) physical processes such as star formation and AGN feedback, which act to lower the total gas mass within r_{500}.

The simulated clusters lie systematically above the mean relation of Chiu et al. (2018) but are in good agreement with the Gonzalez et al. (2013) clusters at low-

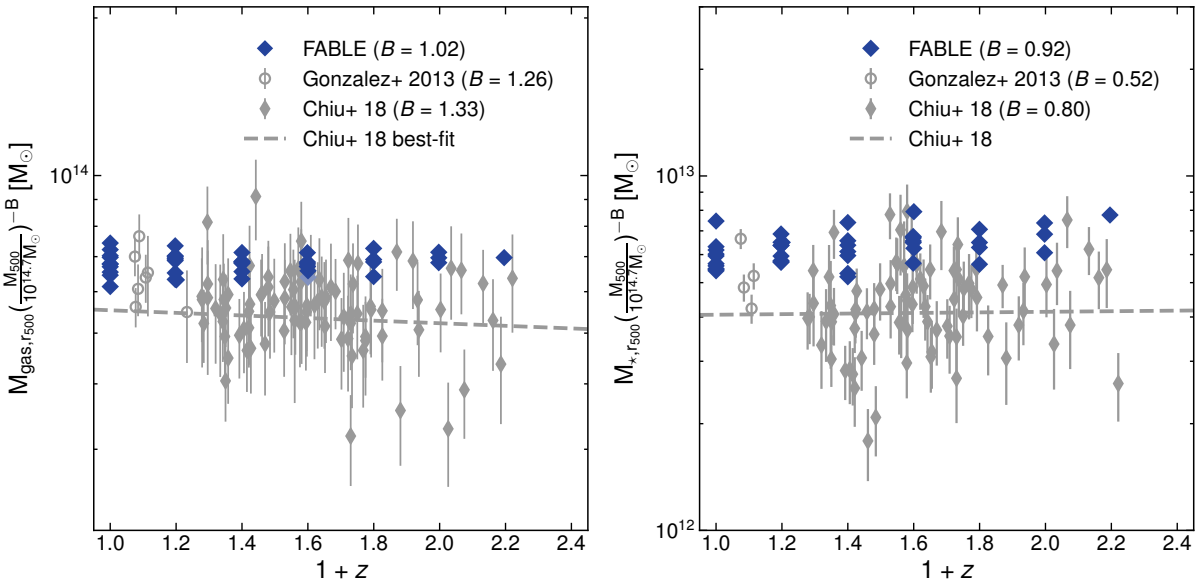


Figure 4.2: The redshift trend of the total gas mass (left) and total stellar mass (right) within r_{500} with respect to the pivot mass $4.8 \times 10^{14} M_\odot$ in comparison to results from Chiu et al. (2018) for a sample of 91 SPT-selected clusters (grey diamonds). We normalise to the pivot mass using the best-fitting mass slope of each sample as indicated in the legend. We compare to FABLE clusters above a mass threshold of $3 \times 10^{14} M_\odot$ at each redshift in order to roughly match the SZ-selected Chiu et al. (2018) sample for which the minimum total mass is roughly constant at this value for all redshifts shown. We additionally plot low-redshift data from Gonzalez et al. (2013) for clusters within the same mass range (open circles). Three of the Gonzalez et al. (2013) clusters do not have total stellar mass measurements.

redshift ($z \lesssim 0.2$). This is understandable given that the Gonzalez et al. (2013) data are based on X-ray hydrostatic masses, whereas the Chiu et al. (2018) mass estimates are effectively weak lensing-calibrated. As discussed in Section 4.1.1 and in previous chapters, weak lensing-based constraints imply that FABLE clusters have too high gas mass at fixed total mass, consistent with the offset seen here.

By extending the mass range of our sample to include low-mass clusters with $M_{500} > 10^{14} M_\odot$ we find that the median gas mass falls slowly with decreasing redshift. This is the approximate lower mass limit of ongoing and future SZ-selected surveys such as SPT-3G (Benson et al., 2014) and CMB-S4 (Abazajian et al., 2016), which suggests that we may be able to observe such a redshift trend in the near future. Qualitatively this agrees with the study by Lin et al. (2012) who find that the gas mass at fixed total mass increases with increasing redshift as $(1+z)^{0.41 \pm 0.14}$ at $z < 0.6$ for a sample spanning a wide mass range ($8 \times 10^{13} \lesssim M_{500} \lesssim 2 \times 10^{15} M_\odot$). Note that this modifies point (2) above as it implies that the low-mass haloes accreted by massive clusters had somewhat larger gas

mass fractions at higher redshift.

The change in gas mass with redshift in low-mass clusters reflects the rapid decrease in the normalisation of the gas mass–total mass relation at $z \lesssim 1$ shown in Fig. 3.3. We attribute this evolution to the increasing effectiveness of gas expulsion by AGN feedback with decreasing redshift due to a corresponding fall in the density (and therefore the binding energy) of haloes of fixed mass. A similar trend was found in the recent simulation studies of Barnes et al. (2017a) and Truong et al. (2018). With this in mind, the observation that clusters with $M_{500} > 3 \times 10^{14} M_{\odot}$ do not show any significant redshift evolution in gas mass (Fig. 4.2) suggests that AGN feedback is ineffective in such massive systems. This is supported by Fig. 4.1, which shows that the gas mass fractions are almost constant on these mass scales, close to the cosmic baryon fraction.

The scatter in total gas mass at fixed redshift appears somewhat smaller in the simulations than the observations. Accounting for the scatter due to measurement uncertainties, Chiu et al. (2018) measure an intrinsic scatter in the gas mass at fixed total mass of 0.05 ± 0.01 dex. For FABLE clusters in the same mass range ($M_{500} > 3 \times 10^{14} M_{\odot}$), we measure an intrinsic scatter of $0.03^{+0.01}_{-0.01}$ dex at $z = 0$ or $0.02^{+0.01}_{-0.01}$ dex at $z = 0.6$. This is slightly lower than the observations, although the scatter is likely biased low due to our small sample size.

4.1.3.2 *Stellar mass redshift trend*

Chiu et al. (2018) find that the total stellar mass at fixed cluster mass varies with redshift as $(1+z)^{0.05 \pm 0.25}$ (dashed line in the right-hand panel of Fig. 4.2). The measured redshift evolution is thus statistically consistent with zero evolution, albeit with large uncertainty. This implies that the total stellar mass at fixed halo mass remains roughly constant with redshift at $z \lesssim 1.25$ for clusters with $M_{500} \gtrsim 3 \times 10^{14} M_{\odot}$. This is consistent with Lin et al. (2012) and Lin et al. (2017), who find no evidence for redshift evolution in the stellar mass–total mass relation at $z \lesssim 0.6$ and $z \lesssim 1$, respectively. In agreement with the observational results, Fig. 4.2 shows no significant change in stellar mass with redshift for FABLE clusters with mass $M_{500} \gtrsim 3 \times 10^{14} M_{\odot}$.

Naively one might expect that massive clusters form predominantly by the accumulation of lower mass haloes. However, this does not seem to be compatible with the shallow slope of the stellar mass–total mass relation, which corresponds to lower mass haloes having higher stellar mass fractions (see Section 4.1.2 and also Fig. 2.4). As such, when a cluster accretes a lower mass halo its total stellar mass fraction is expected to increase. Chiu et al. (2018) hypothesise that the accretion of lower mass haloes must therefore be balanced by a substantial infall

of material from the surrounding environment, which has a significantly lower stellar mass fraction than the cluster itself, particularly at high redshift. Infall from these regions therefore counteracts the increase in the stellar mass fraction that would result purely from accretion of smaller objects.

There are a number of elements that are missing from this toy model however. For example, while it seems difficult to assemble massive clusters from low-mass clusters and groups that have higher stellar mass fractions within r_{500}, clusters will also accrete mass belonging to these haloes that lies outside r_{500} where the stellar mass fraction is smaller. Furthermore, when extending the lower mass limit of the simulated sample below $M_{500} \sim 10^{14} M_{\odot}$ we see a mild increase in the median stellar mass with decreasing redshift. This implies that the stellar mass fractions of galaxy groups were smaller in the past, which goes some way to explaining the relatively low stellar mass fractions of massive clusters at the present day. Cosmological hydrodynamical simulations are in a unique position to quantify the importance of these different processes in a self-consistent manner. To this end, in a future study we plan to track, as a function of time, the proportion of stellar material in FABLE clusters originating from the accretion of low-density material from outside the cluster, the accumulation of lower mass haloes, and in-situ star formation.

The FABLE clusters tend to lie on the upper end of the scatter in the SPT clusters. This appears to be at odds with the $z \approx 0$ comparison in Fig. 4.1 where we demonstrate a good match to observations of the total stellar mass at fixed halo mass. As for the gas mass comparison in the previous section, this offset may be the result of X-ray hydrostatic mass bias in the comparison samples. Indeed, as we discussed in Section 4.1.2, the weak lensing-based analyses of Huang et al. (2018) imply a somewhat lower stellar mass at $M_{500} \sim 3 \times 10^{14} M_{\odot}$ compared with measurements from Gonzalez et al. (2013) based on X-ray hydrostatic masses. On the other hand, the stellar masses derived in Chiu et al. (2018) are likely underestimated given that the limited depth of their imaging data prohibits a measurement of the ICL. Indeed, Chiu et al. (2018) compare their stellar mass–total mass relation to a number of studies in the literature and find a fairly large variation in their normalisations (~ 40 per cent), even after accounting for differences in the assumed initial mass function (IMF) and the method used to estimate the halo mass. Other sources of systematic error that may explain this variation include sample selection biases (e.g. Decker et al. 2019), background subtraction (e.g. Bernardi et al. 2007, 2013; Von Der Linden et al. 2007; van der Burg et al. 2014), or the conversion from luminosity to stellar mass (e.g. Conroy et al. 2009; Bernardi et al. 2017).

The scatter in stellar mass does not change significantly with redshift in either

the observed or simulated samples, albeit with relatively small sample sizes in both cases. Chiu et al. (2018) obtain a best-fitting value for the intrinsic scatter of 0.10 ± 0.01 dex at fixed total mass. This is still larger than the intrinsic scatter in the simulated sample (e.g. $0.06^{+0.02}_{-0.01}$ dex at $z = 0$) but is similar to the intrinsic scatter of 0.11 ± 0.03 dex found in Kravtsov et al. (2018) for clusters at $z \approx 0$, which provides further evidence for a lack of redshift evolution in the scatter. It is encouraging that the intrinsic scatter measured in local samples of clusters, which are often selected on observational signatures closely related to the stellar mass, agrees well with the nearly mass-limited, SZ-selected sample of Chiu et al. (2018). For the latter, clusters are selected independently of their stellar content and should therefore represent an unbiased sampling of the stellar mass fractions in the full cluster population. This lends support to the notion that recent simulations, including FABLE and IllustrisTNG, slightly underestimate the intrinsic scatter in stellar mass of similar mass clusters (see Section 4.1.2).

4.2 BRIGHTEST CLUSTER GALAXIES

BCGs encode important information about the history of their host clusters. For example, in the hierarchical structure formation scenario, cluster mergers typically result in the mergers of their BCGs. As a result, the BCG stellar mass is expected to correlate with the mass of the parent halo (e.g. White & Rees 1978). Indeed, a number of studies have shown that BCG luminosity or stellar mass tends to increase with total cluster mass, albeit with significant scatter (e.g. Lin & Mohr 2004; Gonzalez et al. 2013; Lin et al. 2017; Kravtsov et al. 2018). Furthermore, studies focused on the redshift evolution of BCG stellar mass can provide a great deal of information about the dominant mechanisms responsible for the growth of BCGs and their host clusters, such as minor mergers versus in-situ star formation (Burke & Collins, 2013; Lin et al., 2013; Gozaliasl et al., 2016, 2018; Cooke et al., 2018).

With this in mind, in Section 4.2.2 we study the relationship between BCG mass and cluster mass in the FABLE simulations with comparison to observed BCGs and other recent simulation studies. Then in Section 4.2.3 we investigate the redshift evolution of the BCG stellar mass in FABLE clusters and in Section 4.2.4 we study the BCG stellar mass profiles at different redshifts. Firstly, we summarise the data used for our comparisons to observations.

4.2.1 *Observational data*

In recent years a number of observational studies have constructed large samples of BCGs spanning a wide range in mass and redshift with the aim of better understanding their growth history. We make use of several such studies in the following sections, specifically Lidman et al. (2012), Bellstedt et al. (2016), Zhang et al. (2016), Kravtsov et al. (2018) and DeMaio et al. (2018).

Lidman et al. (2012) and Bellstedt et al. (2016) both construct large samples of BCGs over a wide redshift range with new, published and archived data from various instruments. The Lidman et al. (2012) sample contains 140 BCGs with host cluster mass estimates in the range $2.2 \times 10^{13}\,M_{\odot} \lesssim M_{500} \lesssim 3.7 \times 10^{15}\,M_{\odot}$ and redshifts $0.03 < z < 1.63$. The Bellstedt et al. (2016) sample contains 132 BCGs (102 of which have host cluster mass estimates) spanning the redshift range $0.03 < z < 1.07$ and cluster mass range $1.3 \times 10^{14}\,M_{\odot} \lesssim M_{500} \lesssim 2.3 \times 10^{15}\,M_{\odot}$. We have converted the cluster mass estimates reported in these studies from M_{200} to M_{500} by a factor of 0.72 appropriate for a Navarro-Frenk-White profile (Navarro et al., 1997) with a concentration parameter of $c = 5$, which was performed in the reverse sense in Bellstedt et al. (2016). Masses are estimated from scaling relations between X-ray hydrostatic mass and X-ray luminosity, temperature or gas mass. For a small subset of clusters Lidman et al. (2012) estimate cluster mass from the line-of-sight velocity dispersion. Both studies use the `MAG_AUTO` magnitude estimate from SExtractor to estimate total BCG luminosities and stellar masses. We assume that the `MAG_AUTO` aperture is comparable to a fixed 30 physical kpc radius aperture as suggested in Zhang et al. (2016) (see their appendix B4). We caution, however, that `MAG_AUTO` uses an adaptively scaled aperture, which may introduce additional system-to-system scatter. Individual uncertainties on their stellar mass measurements are not quoted but are expected to be on the order of ≈ 20 per cent (Bellstedt et al., 2016).

Zhang et al. (2016) investigate BCG stellar mass growth using a sample of 106 X-ray selected groups and clusters at $0.07 < z < 1.26$ with deep Dark Energy Survey Science Verification (DES SV) data (Sánchez et al., 2014). Cluster masses are estimated from the weak lensing calibrated mass–temperature relation of Kettula et al. (2013) and BCG stellar masses are measured in circular apertures with physical radii of 32 kpc and 50 kpc. In the following we assume that their 32 kpc radius aperture is equivalent to a 30 kpc aperture. We convert M_{200} to M_{500} by the factor 0.72 as mentioned above. Zhang et al. (2016) derive a redshift-dependent BCG mass–total mass relation for each aperture, which we plot as a dashed line in the following figures.

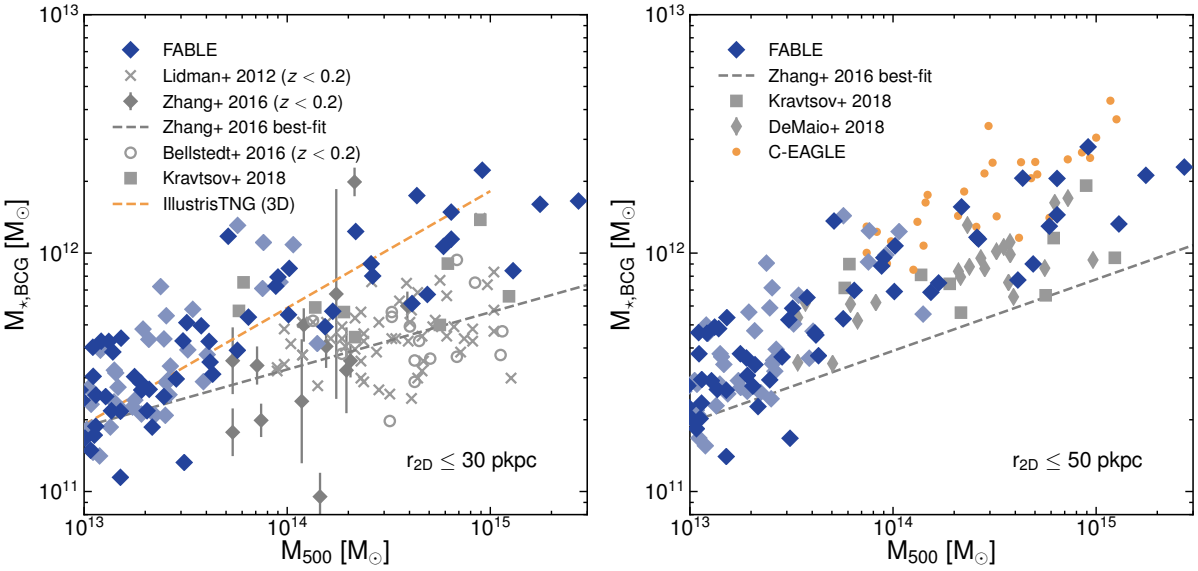


Figure 4.3: The stellar mass of FABLE BCGs at $z = 0$ (diamonds) measured within a projected aperture of radius 30 kpc (left) and 50 kpc (right) as a function of total mass in comparison to observational data (grey points), C-EAGLE clusters (orange circles) and IllustrisTNG (dashed line). For clarity we neglect to show error bars on the total mass estimates. The best-fitting relation from IllustrisTNG is based on stellar masses measured within a spherical rather than a circular aperture. All stellar masses are based on a Chabrier (2003) IMF.

Kravtsov et al. (2018) analyse SDSS data for nine nearby ($z < 0.1$) clusters and derive BCG stellar masses within both a 30 kpc and 50 kpc radius aperture. Three of the nine clusters have direct X-ray hydrostatic mass measurements. Cluster masses for the remaining six are estimated from the mass proxy $Y_{\rm X}$ (Kravtsov et al., 2006) using the scaling relation derived in Vikhlinin et al. (2009) based on X-ray hydrostatic masses. Kravtsov et al. (2018) do not quote individual errors on their stellar mass measurements, although the typical error due to uncertainties in the estimate of the background is small ($\lesssim 10$ per cent).

DeMaio et al. (2018) study 23 galaxy groups and clusters at $0.29 \leq z \leq 0.89$ with *Hubble Space Telescope* imaging. BCG stellar masses are derived within a 50 kpc radius and cluster masses are estimated from a scaling relation between X-ray temperature and X-ray hydrostatic mass (Vikhlinin et al., 2009).

All of these studies assume the Chabrier (2003) IMF in the conversion of luminosities to stellar masses. The same IMF was assumed in the FABLE galaxy formation model.

4.2.2 *BCG stellar mass to total mass relation*

Figure 4.3 shows the stellar mass of FABLE BCGs as a function of their host cluster mass at $z = 0$ in comparison to data from the aforementioned studies, as well as the C-EAGLE and IllustrisTNG simulations. We calculate stellar masses within a 2D radius of 30 kpc (left-hand panel) and 50 kpc (right-hand panel) integrated through the entire simulation volume, which mimics observed BCG luminosities measured within a circular aperture. For the Lidman et al. (2012), Bellstedt et al. (2016) and Zhang et al. (2016) samples we restrict our comparison to low-redshift BCGs at $z < 0.2$.[2] We compare to the full sample of BCGs from DeMaio et al. (2018) but caution that these are situated at somewhat higher redshifts ($0.29 \leq z \leq 0.89$).

The mean FABLE relation is systematically higher than the observed relations by ~ 0.2–0.3 dex, although there is some overlap in the scatter. Still, it is worth pointing out that the FABLE BCGs are much closer to the data than many simulations have been in the past, particularly those without AGN feedback (see e.g. Puchwein et al. 2010, Ragone-Figueroa et al. 2013 and Martizzi et al. 2014). The discrepancy with observations is smallest in comparison to the Kravtsov et al. (2018) and DeMaio et al. (2018) constraints, which measure slightly larger BCG masses at fixed cluster mass than Lidman et al. (2012), Bellstedt et al. (2016) and Zhang et al. (2016). The difference between studies is most likely due to the choice of mass-to-light ratio, which can introduce systematic uncertainties of ~ 0.1–0.2 dex in the stellar mass measurements (see e.g. Conroy 2013). This is not sufficient to explain the discrepancy with the simulations however, which we shall discuss in more detail below. Despite the offset in the normalisation, the slope of the simulated relation appears to be in good agreement with observations in the cluster regime, with signs of a steepening of the relation at group scales. To test this we fit a power-law relation to FABLE haloes with $M_{500} > 6 \times 10^{13} M_{\odot}$, which is the approximate lower mass limit of the data at $z < 0.2$. This yields best-fitting slopes of $0.27^{+0.07}_{-0.07}$ and $0.30^{+0.05}_{-0.05}$ for the 30 and 50 kpc apertures, respectively. These are in good agreement with the values of 0.24 ± 0.08 and 0.30 ± 0.08 derived in Zhang et al. (2016) for their full sample. Kravtsov et al. (2018) measure a slope of 0.39 ± 0.17 using 'total' BCG stellar masses derived from triple Sérsic fits to the light profiles, which is also consistent with our results. On the other hand, Bellstedt et al. (2016) and Lidman et al. (2012) derive significantly steeper slopes of 0.64 ± 0.03 and ≈ 0.63, respectively. This difference may be due to (possibly redshift-dependent) selection effects, as noted in Lidman et al. (2012). It may

[2] For Zhang et al. (2016) individual stellar mass measurements are available only for the 32 kpc aperture.

also indicate a redshift-dependent slope, however we lack a sufficient number of high-redshift clusters in order to constrain this possibility from the simulations. In addition, we predict a substantial scatter in BCG mass at fixed total mass similar to the observations. The intrinsic scatter about the best-fitting relation is $0.15^{+0.02}_{-0.01}$ dex and $0.14^{+0.02}_{-0.01}$ dex for the 30 kpc and 50 kpc apertures, respectively, which is comparable to the values of 0.18 ± 0.02 and 0.19 ± 0.02 derived in Zhang et al. (2016) and the scatter of 0.21 ± 0.09 measured in Kravtsov et al. (2018) for their 'total' BCG masses.

The high stellar masses of FABLE BCGs may reflect a slight overestimate in the abundance of massive galaxies with $\gtrsim 10^{11} M_\odot$ compared with observations of field galaxies at $z \approx 0$ (see Fig. 2.2). As such, further improvements to the modelling of AGN feedback may aid in reducing the stellar masses of the simulated BCGs. Part of the discrepancy may also result from excess in-situ star formation. To test this hypothesis we have calculated the total star formation rate (SFR) within twice the stellar half-mass radius of each BCG. Of the 12 FABLE BCGs in clusters with $M_{500} > 2 \times 10^{14} M_\odot$ at $z = 0$, only one BCG can be considered highly star forming ($\mathrm{SFR} > 10\, M_\odot\, \mathrm{yr}^{-1}$), consistent with the small fraction ($\sim$ 1–5 per cent) of observed BCGs at $z \lesssim 0.1$ (see McDonald et al. 2016 and references therein). Of the remaining 11 BCGs, 6 have a moderate SFR greater than $1\, M_\odot\, \mathrm{yr}^{-1}$. This fraction is somewhat high compared with observed BCGs (e.g. Hoffer et al. 2012; Fraser-McKelvie et al. 2014; Donahue et al. 2015; Fogarty et al. 2015), although it is difficult to make an exact comparison due to large uncertainties in the measured SFRs. We have also compared the SFR as a function of redshift to the measurements of McDonald et al. (2016) for 90 SZ-selected BCGs at $0.25 < z < 1.25$ and find that the simulated BCGs lie on the upper end of the scatter in the data at these redshifts (not shown). Thus, in-situ star formation may be responsible for some of the stellar mass growth in FABLE BCGs, at least at $z \lesssim 1$ (see Section 4.2.3).

One of the dominant mechanisms by which BCGs are expected to grow in mass is via "galactic cannibalism". This describes the process in which smaller satellite galaxies sink towards the cluster centre via dynamical friction, eventually merging with the BCG. Along the way, a significant mass of stars can accumulate in the BCG or the ICL after being stripped from infalling satellites via tidal interactions within the cluster potential. Our results suggest that this process occurs at too high a rate in FABLE clusters, which may explain their overly massive BCGs. Indeed, the total stellar mass in satellites in FABLE clusters is underestimated compared with observations, by as much as $\approx 10^{12} M_\odot$ at $M_{500} \sim 10^{14} M_\odot$ and $\approx 3 \times 10^{12} M_\odot$ at $M_{500} \sim 10^{15} M_\odot$ (not shown). This is consistent with a picture in which the satellites are aggressively stripped of their stars as they move through

the cluster. We expect that this is a consequence of their sizes, which are too large compared with observations. This is because the galaxy size–mass relation in FABLE is relatively unchanged from the original Illustris galaxy formation model, which produced too-large galaxies for a given stellar mass (see e.g. Appendix A in Furlong et al. 2017). Since satellites are preferentially stripped of stars on the outskirts of the galaxy where the gravitational binding strength is weakest (see e.g. Puchwein et al. 2010), larger satellites galaxies of a given stellar mass will lose a larger fraction of their mass. This is consistent with the work of Wang et al. (2019) who find that satellite galaxies in Illustris can lose a significant mass of stars as they approach the central galaxy, in contrast to the predictions of a semi-analytic model. Currently the most viable solution to this problem is to calibrate the galaxy formation model – in particular stellar feedback – to reproduce the observed size–mass relation of galaxies, as was done for the C-EAGLE (EAGLE) model (Crain et al., 2015) and IllustrisTNG (Pillepich et al., 2018a).

Even so, the IllustrisTNG and C-EAGLE clusters also possess significantly more massive BCGs than observed (orange dashed line and orange circles in Fig. 4.3, respectively). We point out that the best-fitting IllustrisTNG relation shown in Fig. 4.3, which is based on stellar masses measured in a spherical 3D aperture, would shift to a slightly higher normalisation for stellar masses measured in projection (by a factor ~ 1.2 on average for FABLE BCGs). In IllustrisTNG the total stellar mass locked in satellites is about 30 per cent lower than observational constraints for a given halo mass (Pillepich et al., 2018b), which could signal excess tidal stripping of satellite galaxies in spite of their realistic sizes (Pillepich et al., 2018a). In contrast, the stellar mass function of satellite galaxies in C-EAGLE is in excellent agreement with a number of observational constraints (Bahé et al., 2017). The cause of the overly massive BCGs in C-EAGLE and IllustrisTNG thus remains unclear. For C-EAGLE, Bahé et al. (2017) find that most of the stellar mass in BCGs is already in place at $z \sim 1$, which implies that further work is needed in modelling their high-redshift progenitors.

It is likely that the FABLE results would compare more favourably with observations if the model were calibrated to produce more realistic galaxy sizes. Even so, the results of C-EAGLE and IllustrisTNG suggest that matching the observed mass and size distribution of galaxies in the field is not, on its own, sufficient to reproduce the observed partitioning of stellar mass into satellites, the BCG and the ICL in galaxy clusters. The fact that BCG masses in FABLE are in similar or better agreement with observations than C-EAGLE or Illustris-TNG in spite of the discrepancy in satellite mass is encouraging.

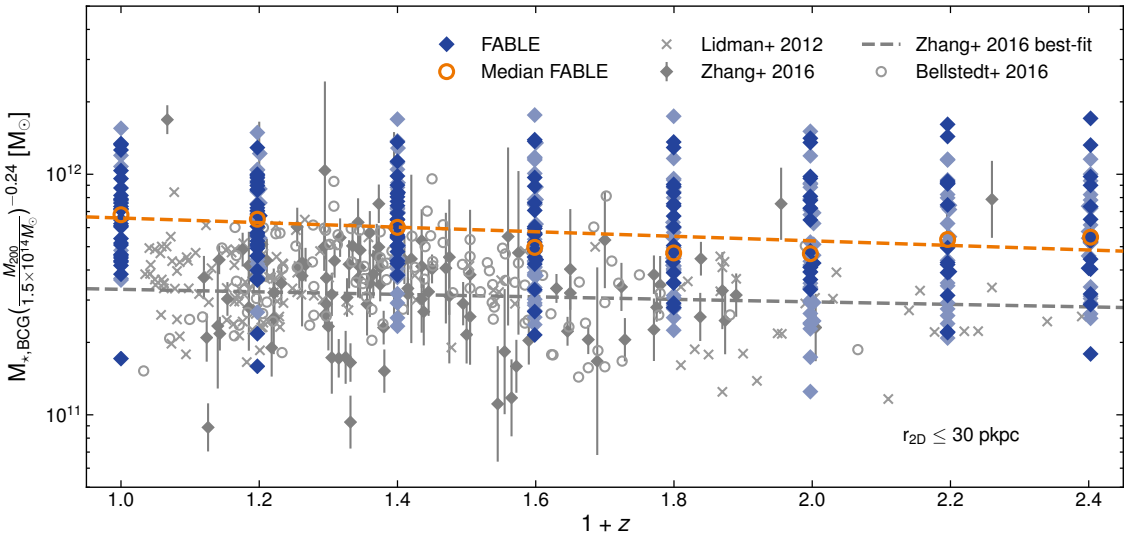


Figure 4.4: The stellar mass of FABLE BCGs (diamonds) as a function of redshift compared to observational data from Lidman et al. (2012) (grey crosses), Zhang et al. (2016) (grey diamonds) and Bellstedt et al. (2016) (grey circles). At each redshift we plot the BCG stellar mass for all FABLE haloes with $M_{200} > 5 \times 10^{13} M_\odot$ to approximately match that of the Zhang et al. (2016) sample. Orange circles indicate the median BCG mass at each redshift and the orange dashed line shows the best-fitting power-law relation to these points in analogy with the best-fitting relation from Zhang et al. (2016) (dashed grey line). The BCG stellar mass is measured within a projected aperture of radius 30 kpc to approximately match the 32 kpc radius aperture of Zhang et al. (2016) and the MAG_AUTO apertures of Lidman et al. (2012) and Bellstedt et al. (2016) (see text). We account for the total mass dependence of the BCG stellar mass with a total mass dependent scaling factor based on the relation of Zhang et al. (2016). We do not take into account the uncertainty in the total mass estimates in this scaling.

4.2.3 *BCG stellar mass redshift trend*

In Fig. 4.4 we plot the stellar mass of FABLE BCGs as a function of redshift in comparison to Lidman et al. (2012), Bellstedt et al. (2016) and Zhang et al. (2016). As motivated in Section 4.2.2, we compare to these data using BCG stellar masses measured within a 30 kpc radius aperture. For our comparison sample we include, at each redshift, all FABLE BCGs in clusters with halo mass $M_{500} > 5 \times 10^{13} M_\odot$, which is the approximate lower halo mass limit of the Zhang et al. (2016) sample. We factor out the total mass dependence of the BCG stellar mass using the BCG–total mass relation derived in Zhang et al. (2016), $M_{\star,\mathrm{BCG}} \propto (\frac{M_{200}}{1.5\times10^{14}M_\odot})^{0.24\pm0.08}(1+z)^{-0.19\pm0.34}$, which has a mass slope similar to that of the FABLE relation for the halo mass range under consideration. As such, Fig. 4.4 represents the redshift trend of the BCG stellar mass at a fixed cluster mass of $M_{200} = 1.5 \times 10^{14} M_\odot$.

The median stellar mass of the FABLE BCG sample (open circles in Fig. 4.4)

is systematically high compared to the data, similar to the $z \approx 0$ comparison in Fig. 4.3. The fact that the offset persists to high redshift implies that a significant proportion of the stellar mass build-up that is responsible for the high masses of FABLE BCGs at fixed cluster mass occurs at $z \gtrsim 1.5$, similarly to the conclusions of Bahé et al. (2017) for the C-EAGLE simulations. It remains unclear whether this results from an excess of in-situ star formation in the main progenitors, an excess of stellar mass in accreted galaxies at high-redshift or the excess tidal stripping of satellite galaxies that manifests in under massive satellites at $z = 0$. In support of the former explanation we find a steady increase in the average BCG SFR with increasing redshift. On the other hand, the redshift trend is in good agreement with that found in the SZ-selected sample of McDonald et al. (2016) at $0.25 < z < 1.25$, with no significant redshift evolution at higher redshifts. It seems unlikely that BCGs are accreting overly-massive satellite galaxies given that the galaxy stellar mass function is in good agreement with observations of field galaxies at $z > 1$ (see Fig 2.3). The exception is an apparent excess of galaxies at $M_\star \lesssim 10^{10} M_\odot$, however these are expected to contribute only a small fraction ($\lesssim 20$ per cent) of the final stellar mass of the BCG (see e.g. fig. 4 of De Lucia & Blaizot 2007). The possibility of excessive tidal stripping is supported by the analysis of a series of cluster zoom-in simulations reported in Puchwein et al. (2010), wherein stars in the ICL at $z = 0$ are preferentially stripped from massive galaxies that fall into the forming cluster at $z > 1$. However, it is unclear to what extent this process can introduce extra stellar mass into the BCG (e.g. within 30 kpc) rather than depositing it into the ICL. We aim to investigate this process in a future, dedicated study of the assembly of FABLE BCGs and ICL similar to that presented in Puchwein et al. (2010), with the inclusion of merger trees that describe the merger history of cluster satellites and the BCG.

The median mass of the FABLE BCG sample displays a fairly shallow redshift trend. A power-law fit to the median relation (orange dashed line in Fig. 4.4) shows that it scales with redshift as $\approx (1+z)^{-0.35}$. This is slightly steeper than the slope of $(1+z)^{-0.19 \pm 0.34}$ derived by Zhang et al. (2016) but is consistent to within 1-sigma. These redshift trends imply that the average stellar mass of simulated and observed BCGs at fixed cluster mass increases mildly with redshift at $z \lesssim 1$ (e.g. by ~ 30 and ~ 15 per cent from $z = 1$ to $z = 0$ for the simulated and observed samples, respectively).

Because the BCG stellar mass correlates with cluster mass, an estimate of the stellar mass growth of individual BCGs requires taking into account the mass growth of their host clusters. For example, Zhang et al. (2016) estimate the BCG growth rate from their best-fitting BCG mass–cluster mass relation using the evo-

lution history of dark matter haloes in the Millennium simulation (Springel et al., 2005b). Extracting the average mass growth history of a sample of haloes with $M_{200} \approx 6.3 \times 10^{13} M_{\odot}$ at $z \sim 1$ and combining this with their redshift-dependent BCG mass–cluster mass relation, Zhang et al. (2016) estimate a BCG stellar mass growth of ~ 35 per cent from $z = 1$ to $z = 0$. Assuming the same average increase in cluster mass (~ 0.27 dex) and the same BCG mass–cluster mass trend used in Zhang et al. (2016), the estimated stellar mass growth of the FABLE BCG sample is ~ 50 per cent between $z = 1$ and $z = 0$. The increase in the median cluster mass of the simulated sample is slightly lower (~ 0.17 dex), which corresponds to a slightly smaller stellar mass growth of ~ 40 per cent. To obtain a true measure of the stellar mass growth of the simulated BCGs will require carefully constructed galaxy merger trees, which we leave to a future study. Nevertheless, these approximations imply a non-negligible stellar mass growth in the simulated BCGs at $z \lesssim 1$, in contrast to observational studies such as Brown et al. (2008), Whiley et al. (2008), Collins et al. (2009) and Stott et al. (2010) who find no evidence of BCG stellar mass growth since $z \sim 1$. On the other hand, it is unclear whether the simulated BCGs grow as rapidly as suggested by more recent studies such as Lidman et al. (2012), Lin et al. (2013) and Bellstedt et al. (2016) who measure almost a doubling of the stellar mass in BCGs from $z \sim 1$ to $z \sim 0$.

The reason for the lack of consensus in the literature remains unclear. Evolutionary studies are complicated by the correlation between BCG mass and host cluster mass and how the cluster mass varies with time. Different methods for accounting for this dependence may be responsible for some of the inconsistencies. Furthermore, deriving BCG stellar masses from imaging data is not a straightforward problem and inconsistent measurements might explain some of the differences. For example, the choice of measurement aperture determines the sensitivity of the observations for probing different BCG growth pathways, since these do not affect the shapes of BCG light profiles in the same way (e.g. Hopkins et al. 2010). In addition, it is likely that the inferred redshift trend depends on how the chosen aperture scales with the brightness, size and redshift of the galaxy. Some studies, such as Zhang et al. (2016), use a fixed aperture for measuring BCG masses whereas others use an adaptively scaled aperture. For example, Lidman et al. (2012) and Bellstedt et al. (2016) use the adaptive `MAG_AUTO` aperture, which adjusts in size depending on the observed surface brightness with the aim of deriving the best possible estimate of the total BCG luminosity. Lin et al. (2017) find that the typical characteristic radius of the `MAG_AUTO` aperture decreases from 36 kpc to 22 kpc from $z \sim 0.4$ to $z \sim 1.0$. This could mean that Lidman et al. (2012) and Bellstedt et al. (2016) measure on average lower BCG masses at high redshift

compared to Zhang et al. (2016), which could explain some of the difference between their inferred growth rates. Indeed, Zhang et al. (2016) find that using the MAG_AUTO aperture increases the inferred BCG mass growth from $z = 1$ to $z = 0$ by 1-sigma from ~ 35 to ~ 70 per cent, which better matches the level of mass growth inferred by Lidman et al. (2012), Lin et al. (2013) and Bellstedt et al. (2016). As noted above, the redshift evolution of the FABLE BCG sample is statistically consistent with the Zhang et al. (2016) constraint, which is understandable given that we have measured BCG masses within a fixed aperture of similar radius.

4.2.4 *BCG stellar mass profiles*

The distribution of stellar mass within a galaxy and how this evolves with time can greatly inform our understanding of the dominant growth pathways of BCGs (e.g. Bernardi 2009; Ascaso et al. 2011; Bai et al. 2014; Furnell et al. 2018) and massive galaxies in general (e.g. Hopkins et al. 2009, 2010; Wuyts et al. 2010; Sonnenfeld et al. 2013; Zahid et al. 2019), including the relative contribution of stars from dry mergers versus in-situ star formation and the importance of feedback processes. In practice this requires the measurement of the surface brightness profiles of massive galaxies situated at different redshifts to infer something about their evolution.

Observational studies of this type have found conflicting results. For example, Stott et al. (2011) find at most a small increase in the scale size of BCGs between $z \sim 1$ and $z \sim 0.2$ and no evidence for a change in the shape of their light profiles. These results are generally confirmed by Bai et al. (2014) for a sample of BCGs at $0.3 < z < 0.9$ with HST imaging data and a local sample from Gonzalez et al. (2005). Recent results from Furnell et al. (2018) for an X-ray selected sample of 329 clusters at $0.05 < z < 0.3$ also show little evolution in the scale size of BCGs, consistent with Stott et al. (2011). Conversely, Bernardi (2009) find that BCGs at $z \sim 0.25$ are up to 70 per cent smaller than their local counterparts while Ascaso et al. (2011) find an increase in the size of BCGs by a factor of ~ 2 between $z \sim 0.5$ and $z \sim 0$ but no change in the shape of their light profiles. These conflicting results largely reflect the difficulty in measuring accurate sizes for BCGs, which depend sensitively on the profile modelling and measured sky background level (see e.g. discussion in Bai et al. 2014). In this section we study the stellar mass profiles of our simulated BCGs to gain some insight on these issues from our model predictions.

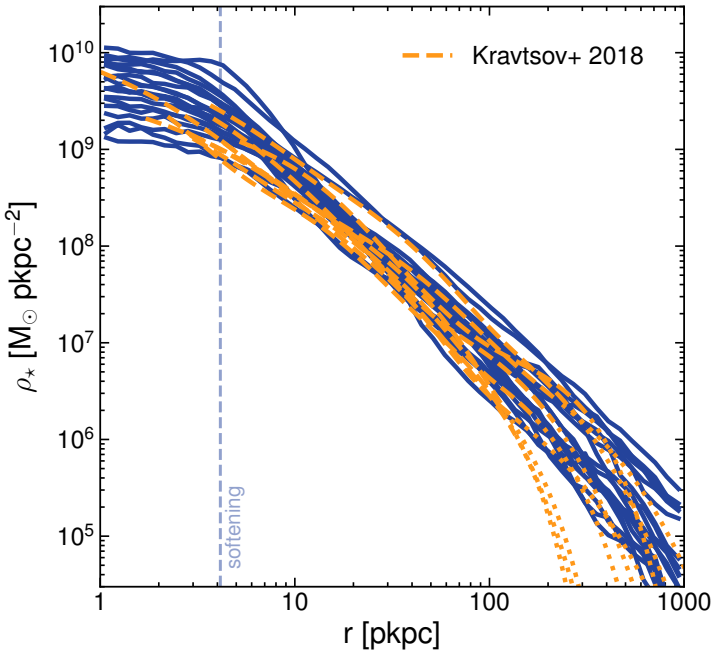


Figure 4.5: Radial profiles of stellar mass surface density for FABLE BCGs (solid lines) at $z = 0$ compared to the best-fitting profiles of observed BCGs from Kravtsov et al. (2018) (dashed lines). Dotted lines indicate extrapolation of the best-fitting triple Sérsic model. The vertical dashed line indicates the gravitational softening length of the simulation below which the simulation predictions are no longer reliable.

4.2.4.1 $z \approx 0$ *comparison with observations*

As a first step we evaluate the consistency of the simulations with observations at $z \approx 0$. In Fig. 4.5 we compare the (projected) stellar mass surface density profiles centred on FABLE BCGs at $z = 0$ with observations of BCGs in local clusters and groups from Kravtsov et al. (2018). The observed profiles are fit with a triple Sérsic model, which are shown in Fig. 4.5 as dashed lines. The observed sample spans the mass range $5.6 \times 10^{13} M_\odot < M_{500} < 1.2 \times 10^{15} M_\odot$ with a median mass of $M_{500} = 2.1 \times 10^{14} M_\odot$. Our comparison sample consists of all $z = 0$ FABLE haloes with $8 \times 10^{13} M_\odot < M_{500} < 1 \times 10^{15} M_\odot$ and has a median mass of $M_{500} = 2.2 \times 10^{14} M_\odot$ similar to the observed sample. In constructing the profiles we include only stars that are gravitationally bound to the main halo, which includes both the BCG and ICL stellar mass but excludes satellites and unbound stars. As mentioned previously, the distinction between BCG and ICL is highly ambiguous, both in simulations and observations. As such, we present the total BCG+ICL stellar mass profiles centred on each BCG.

The FABLE BCGs are slightly too massive at fixed cluster mass compared with the Kravtsov et al. (2018) sample (see Fig. 4.3) and indeed there is a slight offset in

normalisation between the mean simulated and observed profiles. Nevertheless, there is significant overlap between the profiles and excellent agreement in terms of shape across a wide range of scales. We find a similar level of agreement when comparing the three-dimensional simulated profiles with the deprojected stellar mass density profiles derived in Kravtsov et al. (2018). The simulated profiles show an excess of stellar mass in the inner regions ($\lesssim 5$ kpc), however this region accounts for only ~ 20 per cent of the total stellar mass within 30 kpc and thus does not contribute significantly to the overestimated masses of the simulated BCGs seen in Fig. 4.3. Remarkably, the simulated and observed profiles show a very similar level of scatter in stellar mass surface density at fixed radius. Indeed, the intrinsic scatter in the BCG stellar mass–cluster mass relation is in good agreement with Kravtsov et al. (2018) (see Section 4.2.2). This suggests that the sources of intrinsic scatter in observed BCGs, for example feedback processes and merger histories, are adequately modelled in our simulations.

The majority of the extrapolated profiles match the simulated profiles out to several hundred kpc. This suggests that the best-fitting triple Sérsic profile gives a robust measurement of the total stellar mass even beyond the radius of the fit. Indeed, in a recent study of ~ 300 clusters at $z \sim 0.2$, Zhang et al. (2018) find that the BCG+ICL light profile is well described by a triple Sérsic model using high-resolution Dark Energy Survey data. On the other hand, for three of the observed BCGs the extrapolated profiles demonstrate a sharper drop in density at large radii ($\gtrsim 100$ kpc) than predicted by the simulations. These three objects have the smallest extraction radius (the outer radius used in the fitting procedure), which suggests that the triple Sérsic model may underestimate the stellar surface density outside the fitting region if the fitting radius is not sufficiently large ($\gtrsim 150$ kpc in this case).

4.2.4.2 *Redshift evolution*

In Fig. 4.6 we plot radial profiles of stellar mass surface density at redshifts $z = 1$ and $z = 0.2$ in comparison to the stacked BCG profiles of Stott et al. (2011). The Stott et al. (2011) sample includes a high-redshift sample of five BCGs at $0.8 < z < 1.3$ and a low-redshift sample of 19 clusters at $0.15 < z < 0.3$ with deep *Hubble Space Telescope* Advanced Camera for Surveys (*HST*/ACS) imaging data.

Stott et al. (2011) obtain a robust measurement of the typical stellar distribution of BCGs at each epoch by stacking the 1D surface brightness profiles in each sample and fitting the result with a Sérsic profile defined as

$$I(r) = I_e \exp\left\{-b_n\left[\left(\frac{r}{r_e}\right)^{1/n} - 1\right]\right\}, \tag{4.1}$$

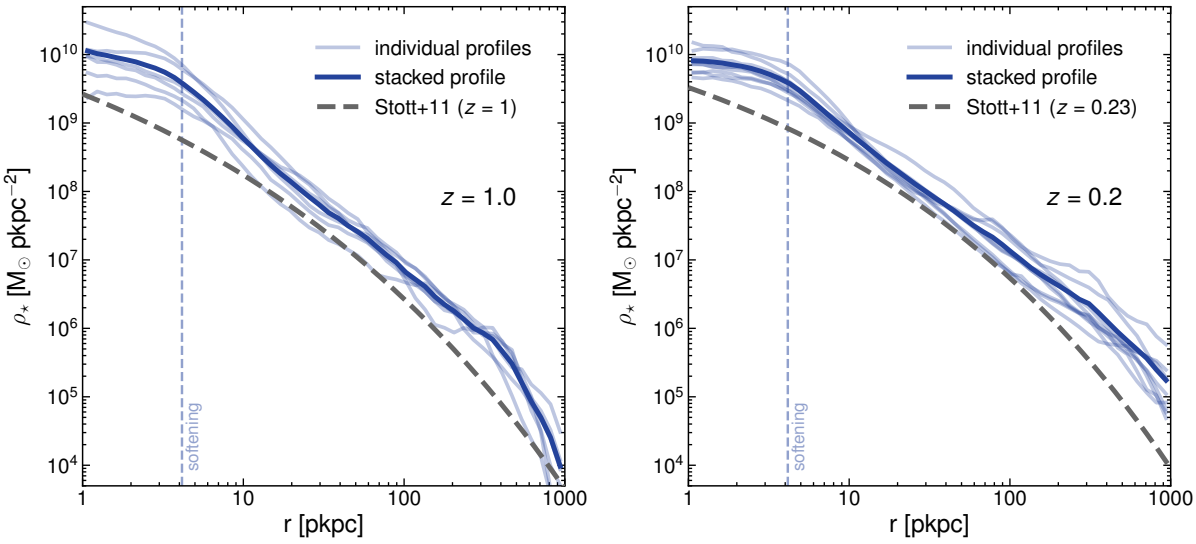


Figure 4.6: Stacked radial profiles of stellar mass surface density for FABLE BCGs (solid lines) compared to the stacked profiles from Stott et al. (2011) (dashed lines) at $z = 1$ (left) and $z \simeq 0.2$ (right). Stott et al. (2011) stack radial surface brightness profiles for a high- and low-redshift sample of BCGs, correcting the profiles to $z = 1$ and $z = 0.23$, respectively. We have converted the best-fitting Sérsic profile for each stack into a stellar mass surface density profile (grey dashed lines) assuming a constant mass-to-light ratio as outlined in the text. The fits shown are those with the Sérsic index allowed to vary during the fit (see text for the best-fitting parameters). Faint blue lines show the individual stellar mass surface density profiles of the FABLE BCGs while the dark blue line shows the stacked profile. The vertical dashed line indicates the gravitational softening length of the simulation.

where $I(r)$ is the surface brightness, r is the radius from the centre of the galaxy, r_e is the effective radius, I_e is the surface brightness at radius r_e, and n is the so-called Sérsic index. The coefficient $b_n = 2n - 0.327$ is chosen such that r_e is the half-light radius defined as the radius that encircles half the light from the galaxy. The profiles are corrected to a common redshift ($z = 1$ and $z = 0.23$ for the high- and low-redshift samples, respectively) using *k* and evolution corrections based on a Bruzual & Charlot (2003) simple stellar population (SSP) model with solar metallicity, formation redshift $z_f = 3$ and a Chabrier (2003) initial mass function. We make use of the Python program EzGal (Mancone & Gonzalez, 2012) to generate Bruzual & Charlot (2003) SSP models with the same parameters in order to convert the best-fitting surface brightness profiles into stellar mass surface density profiles, which are shown as thick dashed lines in Fig. 4.6. The derived solar mass-to-light ratios are 0.88 and 2.71 for the high- and low-redshift samples, respectively.

Weak-lensing mass estimates for clusters in the high-redshift sample range from $M_{200} = 2.9 \times 10^{14} M_\odot$ to $M_{200} = 2.3 \times 10^{15} M_\odot$ (Sereno et al., 2015a). Six

FABLE clusters have $M_{200} \geq 2.9 \times 10^{14} M_{\odot}$ at $z = 1$ and we use this as our high-redshift comparison sample, although we caution that the median halo mass of the sample ($3.5 \times 10^{14} M_{\odot}$) is somewhat lower than that of the observed sample ($6.3 \times 10^{14} M_{\odot}$). The individual profiles are shown as faint blue lines in Fig. 4.6 while the thick blue line shows the stacked profile in analogy with Stott et al. (2011). To construct our low-redshift comparison sample we use the fact that the low-redshift sample of Stott et al. (2011) has an average X-ray temperature similar to that of the high-redshift sample. Since the total mass corresponding to a given X-ray temperature varies with redshift approximately like the self-similar expectation (see Section 3.3.2.2), we scale the lower mass threshold of the high-redshift sample, $M_{200} \geq 2.9 \times 10^{14} M_{\odot}$, by the factor $E(z = 1.0)/E(z = 0.2)$ to obtain an equivalent mass threshold of $M_{200} \geq 4.7 \times 10^{14}\ M_{\odot}$ for our low-redshift ($z = 0.2$) sample. This yields a sample of nine FABLE clusters with a median mass of $M_{200} = 6.9 \times 10^{14}\ M_{\odot}$.

There is a clear offset in normalisation between the simulated and observed stellar mass surface density profiles at both redshifts. The integrated stellar mass within 100 kpc is $\approx$ 2–3 times higher for the simulated stacked profiles, however this is consistent with the overestimate in the BCG stellar masses shown in Fig. 4.3 given the uncertainty in the mass-to-light ratio. Moreover, the shapes of the simulated profiles are consistent with the observations, particularly in the outer regions ($\gtrsim 10$ kpc) where the vast majority of the stellar mass is situated. The profiles deviate at large radii ($\gtrsim 200$ kpc), however Stott et al. (2011) fit the stacked profiles to a surface brightness limit that corresponds to approximately 100 kpc and hence the shape of the profile is not well constrained outside this radius. The simulated profiles show an excess of stellar mass at small radii ($\lesssim 10$ kpc) compared to the observed profiles, similar to the $z \approx 0$ comparison in the previous section.

In analogy with Stott et al. (2011) we fit a Sérsic model of the form of equation 4.1 to our stacked profiles at $z = 1$ and $z = 0.2$ and compare the best-fitting values for the effective radius r_e, which is defined as the half-light radius (or in this case half-mass radius) of the galaxy. For consistency with the observations we fit the model using a least-squares fitting routine (in log-space) out to a radius of 100 kpc, which is approximately the radius at which the observed profiles drop below the surface brightness limit used in the fit. We take the uncertainty in the stacked profile to be the standard deviation of the constituent profiles at each radius.

Since r_e is coupled to the Sérsic index n (Graham et al., 1996), we initially choose a fixed value of $n = 4$ (a de Vaucouleur profile) and compare to the Stott

et al. (2011) results where n was also fixed to this value. The best-fitting effective radii are $r_e = 18.5 \pm 1.7$ kpc and $r_e = 23.6 \pm 2.1$ kpc in physical units at $z = 1$ and $z = 0.2$, respectively. This corresponds to a size increase of 28 ± 3 per cent from high to low redshift, similar to the 35 ± 3 per cent increase found by Stott et al. (2011) who measure $r_e = 32.1 \pm 2.5$ kpc and $r_e = 43.2 \pm 1.0$ kpc for their high- and low-redshift samples, respectively. When excluding the region inside the gravitational softening length, the best-fitting effective radii at $z = 1$ and $z = 0.2$ increase to 22.2 ± 2.9 kpc and 26.0 ± 3.5 kpc, respectively. Note that in this case the size increase from high to low redshift (~ 17 per cent) is even smaller. Performing the fit to larger radii generally biases the best-fitting effective radius towards larger values. This is because the simulated profiles tend to fall off with radius more slowly than assumed in the de Vaucouleur model at $\gtrsim 100$ kpc. For example, when fitting in the radial range 0–500 kpc the best-fitting radii are $r_e = 47.1 \pm 7.0$ kpc and $r_e = 51.2 \pm 9.2$ kpc at $z = 1$ and $z = 0.2$, respectively. In this case the size growth is only 9 ± 2 per cent. Overall these results suggest that a de Vaucouleur profile fit ($n = 4$) to the simulated profiles, as used in many observational studies, leads to an inference of weak size evolution since $z = 1$, consistent with Stott et al. (2011) and in contrast to studies such as Bernardi (2009) and Ascaso et al. (2011).

However, the inferred size evolution is strongly dependent on the choice of model. For example, allowing the Sérsic index to vary (and fitting within 0–100 kpc) we measure $r_e = 21.4 \pm 4.5$ kpc with $n = 7.0 \pm 1.5$ for the $z = 1$ stack and $r_e = 37.9 \pm 10.0$ kpc with $n = 5.7 \pm 1.1$ for the $z = 0.2$ stack. This corresponds to a size increase of 77 ± 26 per cent from high to low redshift, significantly larger than the ~ 28 per cent increase found for the de Vaucouleur profile, albeit with large uncertainties. Stott et al. (2011) also fit their stacked profiles with a free Sérsic model, measuring $r_e = 47.6 \pm 13.7$ kpc with $n = 5.4 \pm 0.9$ at $z \sim 1$ and $r_e = 57.9 \pm 4.5$ kpc with $n = 4.8 \pm 0.2$ at $z \sim 0.2$. Their inferred size growth therefore drops slightly (~ 22 per cent) compared with the $n = 4$ case (~ 35 per cent). The simulated profiles are more sensitive to changes in the Sérsic index because they are, in general, poorly fit by a Sérsic model for reasonable values of $n \lesssim 20$. In particular, outside the fitting radius the simulated profiles show a more gradual change in slope than the free Sérsic fits and are fairly close to a power-law. As a result, the best-fitting Sérsic index diverges as we increase the maximum radius used in the fit. For example, for outer radii greater than ~ 250 kpc the Sérsic index becomes large ($n > 10$), modifying the fit to become closer to a power-law. The same effect has been found in a number of observational studies (e.g. Graham et al. 1996; Gonzalez et al. 2005; Zibetti et al. 2005; Seigar et al. 2007; Stott et al. 2011). These studies find that the outer envelopes of observed

BCGs (which likely contain a significant proportion of ICL) are often poorly fit by a Sérsic profile, biasing the fit to large n. Since the size of this bias depends on the outer radius of the fit, the inferred sizes can be very sensitive to the depth of the data and the background subtraction procedure (e.g. Bernardi et al. 2010). These two factors may explain why a Sérsic model is a good fit to the Stott et al. (2011) profiles but not to the simulations.

The true half-mass radii measured directly from the simulation imply a significant size evolution with redshift that varies depending on the outer radius of integration. For example, for an outer radius of 500 kpc the half-mass radii are $r_e = 39.0$ kpc at $z = 1$ and $r_e = 78.3$ kpc at $z = 0.2$, corresponding to a size increase of ~ 100 per cent. Alternatively, at an integration radius of 100 kpc we find $r_e = 11.2$ kpc and $r_e = 19.5$ kpc, which correspond to a size increase of ~ 74 per cent. This implies significant BCG size growth, with the caveat that we are implicitly assuming that the BCG abruptly ends (and the ICL begins) at a radius of 100 kpc when in reality there is no such distinction. This size growth is consistent with Ascaso et al. (2011) who measure a size increase of 106 ± 63 per cent between $z \sim 0.5$ and $z \sim 0$. Indeed, Ascaso et al. (2011) fit a Sérsic model combined with an exponential component out to ~ 100 kpc, which is a better fit to their observed profiles and likely avoids some of the biases associated with the single Sérsic fit described above.

The size growth inferred from the true half-mass radii increases as we enlarge the integration radius. This indicates that the build-up of stellar mass in BCGs between $z = 1$ and $z = 0.2$ occurs more rapidly at large radii. Indeed, from Fig. 4.6 it is clear that the slope of the simulated profiles at $\gtrsim 100$ kpc is noticeably shallower in the low redshift sample (the surface mass density scales with radius like $\sim r^{-1.7}$ at $z = 0.2$ as opposed to $\sim r^{-2.0}$ at $z = 1$). This implies that, at $z \lesssim 1$, stellar mass growth occurs predominantly in the ICL at large radii. This interpretation is consistent with the relatively mild growth of the BCG stellar mass at $z < 1$ discussed in Section 4.2.3, although we caution that the build-up of ICL in our simulations may be overestimated due to excessive tidal stripping of satellite galaxies as discussed in Section 4.2.2.

Overall our findings confirm those of previous observational studies (e.g. Bernardi et al. 2010, 2013; Bai et al. 2014) which show that the comparison between best-fitting model parameters is ambiguous due to the coupling between parameters (e.g. effective radius and Sérsic index; see e.g. Graham et al. 1996) and systematics in the sky background measurement (and therefore the outer radius to which the BCG profile can be reliably fit). Some of the biases associated with the former can likely be alleviated by a more suitable choice of model, for

example the triple Sérsic profile used in Kravtsov et al. (2018), which we showed in the previous section provides a reasonable description of the BCG profiles out to large radii. However, in that comparison we also found that the profiles with the smallest fitting radii significantly underestimate the stellar mass surface density at large radii, similar to the single Sérsic fits described above. This can considerably bias the inferred scale radius (and potentially the total stellar mass measurement), with important consequences for studies of BCG growth. Our results suggest that careful background subtraction, sufficiently deep data (and/or stacking), and parametric models that accurately describe the outer profiles of BCGs and the surrounding ICL are all required if further progress in the study of BCG growth is to be made. In future, treating the simulations in the same manner as real data using mock observations will help shed further light on these issues.

4.3 SUMMARY

In this work we have studied the total baryon content and BCG properties of galaxy clusters and groups in the FABLE suite of cosmological hydrodynamical simulations. We have investigated the halo mass and redshift dependence of the total gas and stellar mass within r_{500} over a wide halo mass range with comparison to observational constraints and other simulation predictions. We have also studied the stellar masses of the BCGs that form at the centre of these massive haloes, including their correlation with host cluster mass and redshift. In addition, we analysed the stellar mass surface density profiles of the simulated BCGs relative to observations at $z \approx 0$ and $z \approx 1$, highlighting potential biases in observational studies of BCG growth. The main conclusions of this work are summarised in the following points.

- The total gas mass and total stellar mass of FABLE groups and clusters as measured within r_{500} are in excellent agreement with observational constraints based on X-ray hydrostatic mass estimates at $z \approx 0$ across two decades in halo mass ($M_{500} \approx 10^{13} - 10^{15} M_{\odot}$). However, comparison with weak lensing-calibrated constraints implies that the simulated systems are too gas-rich. The level of baryon depletion implied by the weak lensing-based constraints presents a challenge to current models of cluster formation, which highlights the need for improvements in the modelling of AGN feedback and/or the addition of previously neglected physical processes within hydrodynamical simulations.

- For a sample of relatively massive clusters with $M_{500} > 3 \times 10^{14}\,M_{\odot}$ we find that the total gas mass and total stellar mass within r_{500} are approximately independent of redshift at $z \lesssim 1$ in the simulations, in agreement with recent constraints using SZ-selected cluster samples (Chiu et al., 2018). For lower mass samples the simulations predict a non-negligible redshift evolution in these quantities – for low-mass clusters ($M_{500} \lesssim 3 \times 10^{14} M_{\odot}$) in the case of gas and galaxy groups ($M_{500} \lesssim 10^{14} M_{\odot}$) in the case of stellar mass.

- The predicted relation between BCG stellar mass (measured within a radius of 30 and 50 kpc) and host cluster mass has a similar slope and intrinsic scatter to observational constraints at $z \approx 0$, but a somewhat higher normalisation (by ~ 0.2–0.3 dex). This offset is similar to, or less than, that of the recent C-EAGLE and IllustrisTNG simulations. This may result from an overabundance of massive galaxies ($\gtrsim 10^{11} M_{\odot}$) in the field environment, excess in-situ star formation within the BCGs, or excessive tidal stripping of stars from satellite galaxies. Further investigation is required to identify the relative contribution of these different processes to BCG stellar mass growth.

- The average stellar mass of FABLE BCGs at fixed cluster mass increases mildly with decreasing redshift (by ~ 30 per cent from $z = 1$ to $z = 0$) in contrast to observational studies that find no evidence for evolution (e.g. Whiley et al. 2008; Collins et al. 2009; Stott et al. 2010). An accurate measure of the stellar mass growth of individual BCGs within FABLE is left for future study, however an approximate calculation suggests that the stellar mass growth of simulated BCGs between $z = 1$ and $z = 0$ is on the order of ~ 50 per cent, which lies in between recent observational constraints that find mild (~ 35 per cent; Zhang et al. 2016) or fairly rapid (~ 100 per cent; Lidman et al. 2012; Bellstedt et al. 2016) BCG growth.

- Stellar mass surface density profiles centred on FABLE BCGs are in excellent agreement with $z \approx 0$ observations from Kravtsov et al. (2018), which are based on tripe-Sérsic fits to the observed light profile out to ~ 100 kpc. The extrapolation of these fits to larger radii agrees well with the simulated profiles out to several hundred kpc in the majority of cases, although the simulated profiles suggest that the stellar mass density can be underestimated if the outer radius of the fit is not sufficiently large.

- We study the evolution of the simulated profiles with comparison to the stacking analysis of BCG light profiles performed by Stott et al. (2011) for clusters at $z \sim 1$ and $z \sim 0.2$. We find that a single Sérsic model tends to

underestimate the stellar mass density in the simulations beyond the surface brightness limit of the observations ($\gtrsim 100$ kpc) as, on these scales, the simulated profiles follow a close to power-law shape out to several hundred kpc. The slope at large radii is slightly shallower at $z = 0.2$ compared with $z = 1$ such that the inferred size growth is highly sensitive to the outer radius of the fit. This highlights the need for sufficiently deep imaging and careful background subtraction in addition to an appropriate choice for the assumed light profile.

Further study of BCG properties in FABLE will likely benefit from mock observations that allow a like-for-like comparison with the data. Moreover, reconstruction of the merger histories of galaxies in the simulations will enable the study of the evolution of individual BCGs and the fate of infalling satellite galaxies. This will further our understanding of the dominant processes that control the partitioning of stars into satellites, the BCG and the ICL, as well as the relative contribution of accreted haloes to the overall stellar mass and gas mass build-up of massive clusters. Analyses such as these will also help to pinpoint the root causes of the remaining discrepancies with observations in order to highlight the areas that require the most improvement, such as the modelling of AGN feedback.

5 CONCLUSIONS AND FUTURE RESEARCH

5.1 CONCLUSIONS

In this thesis I have employed cosmological hydrodynamical simulations to study a broad selection of astronomical phenomena that arise from the complex physics governing the formation and evolution of galaxy groups and clusters. To this end, I have led the development of the *Feedback Acting on Baryons in Large-scale Environments* (FABLE) project, with the aim of extending and improving previous galaxy formation models to reproduce observations across a vast dynamic range, from low-mass galaxies to galaxy groups and massive clusters.

The FABLE project has met with great success in this endeavour, as I have demonstrated through extensive and careful comparisons with observations throughout this thesis. For example, in Chapters 2 and 4 I showed that the integrated baryon contents of FABLE groups and clusters are in very good agreement with $z \approx 0$ observations across a wide halo mass range. This represents a vast improvement over the original Illustris galaxy formation model on which the FABLE model is based. At the same time, FABLE maintains an excellent match to the observed stellar mass function of galaxies over three decades in stellar mass, including its redshift evolution. The simulated groups and clusters are also in good agreement with a range of observed scaling relations, such as the X-ray luminosity–total mass and Sunyaev-Zel'dovich (SZ) signal–total mass relations, as compared with both low-redshift ($z \approx 0$; see Chapter 2) and high-redshift samples ($z \lesssim 1.0$; Chapter 3). Furthermore, in Chapter 2 I showed that the intracluster medium (ICM) is realistically distributed in galaxy groups and clusters at $z \approx 0$, with deviations arising only in the cluster core regions.

Cosmological hydrodynamical simulations like those in the FABLE project provide an opportunity to investigate aspects of group and cluster formation that are difficult or impossible to constrain with current observations. For example, observational constraints on the redshift evolution of galaxy cluster and (especially) group properties, such as their total baryon content, their central and satellite

galaxies, and the cluster scaling relations, are currently limited by the paucity of known high-redshift clusters and the difficulty in accounting for selection biases in small, often heterogeneous samples. Simulations can provide constraints over a wider mass and redshift range than is accessible by current observations, as well as to aid in understanding sample selection effects. Significant progress has been made in the hydrodynamical simulation of cluster and galaxy formation in recent years that is just now facilitating such studies (e.g. Le Brun et al. 2016; McCarthy et al. 2017; Barnes et al. 2017a; Truong et al. 2018). This is ideal timing given the expected incoming flux of new group and cluster samples from ongoing and upcoming surveys such as eROSITA, SPT-3G and Advanced ACTpol. Detailed follow-up observations with experiments such as the Dark Energy Survey (DES), Euclid, the Large Synoptic Survey Telescope (LSST), the James Webb Space Telescope (JWST; Gardner et al. 2006) and *Athena* will provide precise weak lensing mass estimates, galaxy stellar mass measurements and X-ray measurements of the ICM among other valuable data. Simulations can provide useful predictions and guidance for these upcoming surveys as well as to aid in understanding the underlying cluster physics. To this end, in Chapters 3 and 4 I have utilised the FABLE simulations to make predictions for the redshift evolution of the X-ray and SZ scaling relations, the total gas and stellar content of clusters and the properties of their central galaxies. The main conclusions of this thesis may be summarised as follows.

- The results of FABLE demonstrate that good agreement with observations can be achieved in terms of the global properties of the ICM with a relatively simple model for thermal AGN feedback. However, consistent with the results of several other recent simulation studies, deviations from the observed properties of cluster cores suggest that a more sophisticated modelling of the physics of AGN feedback – or additional physical processes previously neglected in simulations – are required to accurately reproduce the observed thermodynamic profiles of the ICM in the central regions. In particular, many models (including FABLE) struggle to reproduce the observed dichotomy between cool-core and non-cool-core clusters.

- The redshift evolution of the X-ray scaling relations in FABLE show significant deviations from the simple self-similar expectation, in general agreement with other recent simulation studies of this type. In particular, the normalisations of the relations evolve positively with respect to self-similarity due to a combination of factors, such as non-thermal pressure support from bulk motions and turbulence at high redshift and the increased efficiency of AGN

feedback in lower mass haloes at fixed redshift and at lower redshift at fixed halo mass. This has important implications for the use of mass proxies, such as the commonly-used $Y_{\rm X}$ parameter, for estimating the masses of distant clusters. Furthermore, the intrinsic scatter in the relations tends to decrease with increasing redshift and halo mass, which highlights the need for a mass- and redshift-dependent parametrization of the intrinsic scatter in order to properly account for selection biases in observed samples.

- The scaling relation between SZ signal and total mass is in good agreement with low-redshift observations across two decades in halo mass. At higher redshifts up to $z \sim 1$ the simulated clusters are a good match to clusters detected by *Planck* and the SPT-SZ 2500 deg^2 survey. Relatively small differences in the slope, normalisation and redshift evolution of the SZ signal–total mass relation between different simulation and observational constraints have a large impact on the expected number of high-redshift clusters in an SZ survey such as SPT-3G. This emphasises the need for improved modelling of the cluster scaling relations if the potential of future cluster surveys to probe cosmology is to be fully realised.

- The total mass of baryons in FABLE groups and clusters as a function of halo mass is in good agreement with observational constraints based on X-ray hydrostatic cluster mass estimates. On the other hand, constraints based on weak lensing-derived masses imply that FABLE haloes are too gas-rich. A thorough understanding of the X-ray hydrostatic mass bias is required such that reliable comparisons between simulations and observations can be made, in particular with regards to the calibration of sub-grid models such as AGN feedback. The large baryon depletion factor implied by some weak lensing-based constraints provides further reason to revisit the modelling of feedback in hydrodynamical simulations, as well as to investigate potential biases in the weak lensing measurements. The total baryon mass of intermediate-mass clusters ($M_{500} \gtrsim 3 \times 10^{14} M_{\odot}$) is approximately independent of redshift at $z \lesssim 1$, in agreement with the results of recent studies using SZ-selected cluster samples. On the other hand, for lower mass systems the simulations predict a decrease in total gas mass at fixed halo mass with decreasing redshift (see Section 3.3.2.1 and 4.1.3.1). This mass regime ($M_{500} \gtrsim 10^{14} M_{\odot}$) will be accessible to surveys such as SPT-3G and Advanced ACTpol.

- Central brightest cluster galaxies (BCGs) have somewhat high stellar masses,

similar to the recent C-EAGLE and IllustrisTNG simulations. In both FABLE and C-EAGLE an offset persists out to $z \sim 1$. This hints at a shortcoming of current galaxy formation models in the modelling of BCGs, perhaps related to excess in-situ star formation or excessive tidal stripping of stars from cluster satellite galaxies. The average stellar mass growth of simulated BCGs is intermediate between recent observational constraints, which find either slow or fast mass growth at $z \lesssim 1$ (~ 35 to ~ 100 per cent increase in stellar mass from $z = 1$ to $z = 0$). Radial stellar mass profiles centred on the simulated BCGs present a very similar shape to observed profiles at $z \approx 0$. Comparison with observations at $z \sim 0$ and $z \sim 1$ suggests that observational constraints on the stellar mass profiles and their redshift evolution can be biased by the assumed light profile and the outer fitting radius.

5.2 FUTURE RESEARCH

There are many possible avenues of further investigation with the FABLE project, such as the origin and redshift evolution of the X-ray hydrostatic mass bias or the build-up of intracluster stars. In addition, although the FABLE project represents a substantial step forward in the modelling of galaxy and galaxy cluster formation, there are also several areas for improvement. For example, the modelling of AGN feedback or the addition of other physical processes such as anisotropic thermal conduction or cosmic rays. Below I go into more detail about possible ideas for future research that could follow on from this work.

X-ray hydrostatic mass bias

The issue of X-ray hydrostatic mass bias has featured heavily in this thesis because of its importance for the reliable comparison of cluster simulations to observations as a function of halo mass. Therefore, an area of possible future research with FABLE would be a detailed investigation into the sources of X-ray hydrostatic mass bias in simulated clusters and groups. Hydrodynamical simulations are uniquely placed to separate the bias due to lack of hydrostatic equilibrium (e.g. Lau et al. 2009, 2013; Nelson et al. 2014a; Biffi et al. 2016; Shi et al. 2016) from the biases associated with the X-ray measurements, such as projection effects (e.g. Meneghetti et al. 2010), parametric modelling biases (e.g. Rasia et al. 2006; Piffaretti & Valdarnini 2008), spectroscopic temperature bias (e.g. Rasia et al. 2006; Nagai et al. 2007a; Piffaretti & Valdarnini 2008; Meneghetti et al. 2010; Rasia et al. 2012; Henson et al. 2017) or gas inhomogeneities (e.g. Mathiesen & Evrard

2001; Rasia et al. 2012, 2014; Khedekar et al. 2013).

With the FABLE simulations, which are performed with the moving-mesh hydrodynamics scheme AREPO (Springel, 2010b), it will be possible to test the impact of temperature inhomogeneities on X-ray hydrostatic mass estimates complementary to previous studies based on SPH or AMR simulations (e.g. Mathiesen & Evrard 2001; Rasia et al. 2012, 2014). In SPH codes for example, numerical inaccuracies may generate more temperature structure than is present in real clusters (e.g. Sijacki et al. 2012) and may artificially suppress the mixing of low-entropy, high-density gas clumps. In particular it will be possible to test the findings of Henson et al. (2017), who find that the X-ray spectroscopic temperature measurements of MACSIS clusters are significantly biased by cool gas on the cluster outskirts. Another advantage of the FABLE suite is the wide dynamic range of the simulated sample, which will enable constraints on the mass and redshift dependence of the X-ray mass bias. Both of these aspects will be particularly important for the mass calibration of upcoming samples, which will span an increasingly wide range of masses and redshifts.

Understanding the dominant sources of the X-ray hydrostatic bias with such a study may help explain the lack of consensus in previous constraints on the mass bias from both simulations and observations. This may also shed some light on the possible tension between cosmological parameter constraints from cluster counts and *Planck* primary CMB constraints. Such a study will complement upcoming constraints on the mass bias from large, representative cluster samples with mass estimates from weak lensing surveys (e.g. Laureijs et al. 2011; LSST Dark Energy Science Collaboration 2012; Spergel et al. 2015), CMB lensing (e.g. Louis & Alonso 2017; Zubeldia & Challinor 2019) and high-resolution, spatially-resolved X-ray spectroscopy (e.g. Rossetti et al. 2016; Lovisari et al. 2017; Ota et al. 2018).

The fate of infalling satellites and the build-up of intracluster stars

As noted in Chapter 4, one of the shortcomings of the FABLE model is that the total stellar mass in cluster satellites is underestimated compared to observations. A possible cause of this discrepancy is that satellites experience excessive tidal stripping of their stars in the cluster environment as a consequence of their large sizes. This can be tested by tracking the stars that belong to a satellite over its lifetime. Of the 27 zoom-in simulations in the FABLE suite, 14 of these have frequent snapshots spaced equidistant in redshift by $\Delta z = 0.05$ at $z < 3$ and $\Delta z = 0.2$ at $5 < z < 8$. This high time resolution can be exploited to trace the evolution of the simulated galaxies in detail, even for satellites that are rapidly

tidally stripped of stars.

It is expected that many of the stars stripped from satellites will end up in the ICL, with some fraction settling in the BCG. However, this proportion – and its dependence on properties of the satellite, such as its size, mass and time since infall, and of the cluster halo, such as its mass and dynamical state – is not fully understood. Furthermore, it will be possible to determine whether ICL that may be missed in observations could alleviate tensions between constraints on the cosmic baryon fraction and the baryon fractions measured in galaxy clusters, particularly those based on weak lensing masses.

Another interesting route of investigation is the ram pressure stripping of gas from satellite galaxies. In particular, the properties of the hydrodynamics code can be exploited to determine whether previous findings of in-situ star formation in ram-pressure stripped gas were significantly affected by numerical inaccuracies (e.g. Puchwein et al. 2010) – for example, fluid instabilities may be able to disrupt the ram pressure-stripped clouds of gas, which can suppress the star formation happening within.

AGN feedback and additional physics

It is clear from the results of FABLE and several other recent studies (e.g. Barnes et al. 2017b, 2018b) that further improvement in the ability of cosmological hydrodynamical simulations to reproduce the observed thermodynamic profiles of the ICM will require continued development of sub-grid models for AGN feedback and/or the addition of previously neglected physical processes.

The radio-mode of AGN feedback in FABLE employs the same model as Illustris, which mimics the inflation of jet cavities by injecting off-centre, hot bubbles (Sijacki et al., 2007). As an alternative to this model, IllustrisTNG employ a mechanical mode of feedback that mimics a kinetic AGN wind (Weinberger et al., 2017a). This method does not overly heat or dilute the central gas as in Illustris and FABLE, resulting in more realistic thermodynamic profiles (Weinberger et al., 2017a; Barnes et al., 2018b). On the other hand, Barnes et al. (2018b) show that the IllustrisTNG model struggles to produce cool-core and non-cool-core clusters in the observed proportions. Other works have modelled the inflation and evolution of the jet cavities themselves and their impact on cluster haloes (e.g. Dubois et al. 2011; English et al. 2016; Bourne & Sijacki 2017; Weinberger et al. 2017b; Bourne et al. 2019). Despite the success of such models, there is still little consensus on which processes are responsible for the efficient and largely isotropic coupling of the jet energy to the ICM. A number of mechanisms have been proposed, including

shock heating (e.g. Randall et al. 2015; Li et al. 2017), sound waves (e.g. Fabian et al. 2003, 2017; Ruszkowski et al. 2004), turbulence (e.g. Banerjee & Sharma 2014; Zhuravleva et al. 2014), mixing (e.g. Hillel & Soker 2017a,b) and cosmic ray production (e.g. Sijacki et al. 2008; Pfrommer 2013).

Some recent studies have also highlighted the importance of previously neglected physical processes such as anisotropic thermal conduction (Kannan et al., 2016, 2017; Yang & Reynolds, 2016a; Barnes et al., 2018a) and cosmic rays (Simpson et al., 2016; Pakmor et al., 2016; Pfrommer et al., 2017; Ruszkowski et al., 2017). Thermal conduction from the hot outer parts of the ICM to the inner regions may provide some of the heating required to offset radiative cooling losses in cluster cores, reducing the burden on the AGN (e.g. Zakamska & Narayan 2003; Voigt & Fabian 2004; Ruszkowski et al. 2011; Yang & Reynolds 2016a). Furthermore, Kannan et al. (2017) and Barnes et al. (2018a) demonstrate that anisotropic thermal conduction leads to increased mixing that significantly improves the coupling between the AGN feedback energy and the ICM. In addition, observations of synchrotron and inverse Compton emission in X-ray cavities implies a significant non-thermal component in these cavities, such as cosmic rays and magnetic fields. These cosmic rays can escape from the buoyantly rising bubbles and heat the ICM by diffusion or disruption of the bubble by surface instabilities (Guo & Oh, 2008).

Final remarks

An abundance of new multi-wavelength data is expected in the coming years from ongoing surveys with *Chandra*, XMM-Newton, SPT-3G, Advanced ACTpol, DES and the HSC-SSP survey (Aihara et al., 2018) among others (see e.g. Pratt et al. 2019 and references therein). These will be followed in the near future by missions such as LSST, Euclid, JWST, eROSITA, XRISM (Tashiro et al., 2018) and the Simons Observatory (The Simons Observatory Collaboration, 2019) and in the more distant future with *Athena* and proposed experiments such as CMB-S4, AXIS (Mushotzky, 2018) and Lynx (The Lynx Team, 2018). This new data will provide invaluable insight into a broad range of group and cluster phenomena over an unprecedented mass and redshift range. Just as simulations will help guide these surveys and aid the interpretation of the results, the wealth of understanding that stems from this new data will help inform the development of the next generation of simulations. It is this close interaction between ever more precise observations and increasingly sophisticated theoretical models that is vital to build a complete picture of the formation and evolution of galaxy groups and clusters and to further their use as probes of cosmology.

A AGN FEEDBACK MODELS

The FABLE simulations implement a series of physical models based on those of the Illustris galaxy formation simulation. With the Illustris model as our starting point, we have updated the sub-grid model for AGN feedback to reproduce the massive end of the present-day galaxy stellar mass function (GSMF) and the hot gas content of massive haloes with $M_{500} \approx 10^{13} - 10^{14} M_{\odot}$. Multiple variations of AGN feedback were tested, with our preferred model corresponding to the fiducial FABLE model.

Here we present three additional variations of our AGN feedback model and their corresponding impact on the $z = 0$ GSMF and the stellar and gas fractions of massive haloes. Each model has been implemented in a periodic box of length 40 h^{-1} (comoving) Mpc on a side with initial conditions as described in Section 2.2.1. These simulations differ only by the AGN feedback parameters listed in Table A.1. The meaning of each parameter is explained in Section 2.2.4. For reference, we also list the parameters used in Illustris.

In Fig. A.1 we compare the $z = 0$ GSMFs of the four models. The 'stronger

Table A.1: Parameter values for the AGN feedback models presented in this appendix in comparison to that of Illustris. The parameter $\chi_{\rm radio}$ is the fraction of the Eddington accretion rate below which BHs operate in the radio-mode and above which they operate in the quasar-mode; ϵ_r is the radiative efficiency of BH accretion; ϵ_f is the thermal coupling efficiency of the quasar-mode; Δt is the accumulation time period between quasar-mode feedback events in Myr; ϵ_m is the coupling efficiency of the radio-mode and $\delta_{\rm BH}$ is the fractional increase in BH mass required to trigger a radio-mode feedback event.

	$\chi_{\rm radio}$	ϵ_r	ϵ_f	Δt (Myr)	ϵ_m	$\delta_{\rm BH}$
ILLUSTRIS	0.05	0.2	0.05	–	0.35	0.15
WEAK RADIO	0.05	0.1	0.1	–	0.4	0.001
STRONGER RADIO	0.05	0.1	0.1	–	0.8	0.01
QUASAR DUTY CYCLE	0.05	0.1	0.1	25	0.4	0.001
FIDUCIAL	0.01	0.1	0.1	25	0.8	0.01

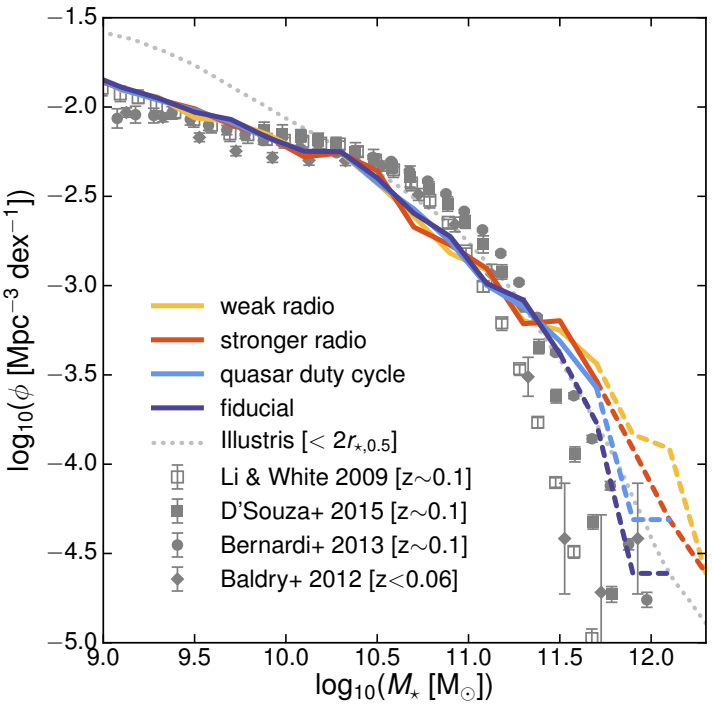


Figure A.1: The galaxy stellar mass function at $z = 0$ for different AGN feedback models (lines) compared to observations (symbols with error bars). Lines become dashed at the high-mass end when there are fewer than 10 objects per 0.2 dex stellar mass bin. Here the stellar mass of a galaxy is defined as the mass of stars bound to the subhalo within twice the stellar half-mass radius. The grey dotted line shows the equivalent stellar mass function in Illustris. The observational data are as shown in Fig. 2.2.

radio' model produces a slightly lower abundance of galaxies at the high mass end compared with the 'weak radio' model. This implies that less frequent but more energetic bubble injections are slightly more efficient at suppressing star formation in massive haloes, partly by displacing gas from the dense central regions and partly by heating the surrounding gas to higher temperatures.

The 'quasar duty cycle' model introduces a quasar-mode duty cycle to the 'weak radio' model, which leads to a significant reduction in the abundance of massive galaxies. This implies that periodic heating is much more effective at suppressing star formation than continuous thermal feedback. Physically, this is because the gas is heated to higher temperatures, reducing cooling losses and slowing the rate at which the gas can condense to form stars. The GSMF of the 'quasar duty cycle' model is similar to Illustris at the high mass end, despite using a far gentler form of radio-mode feedback.

In the left hand panel of Fig. A.2 we plot the median stellar mass fraction as a function of halo mass for each of the models at $z = 0$. The stellar fractions are largely consistent with the difference in the GSMFs between different models. That

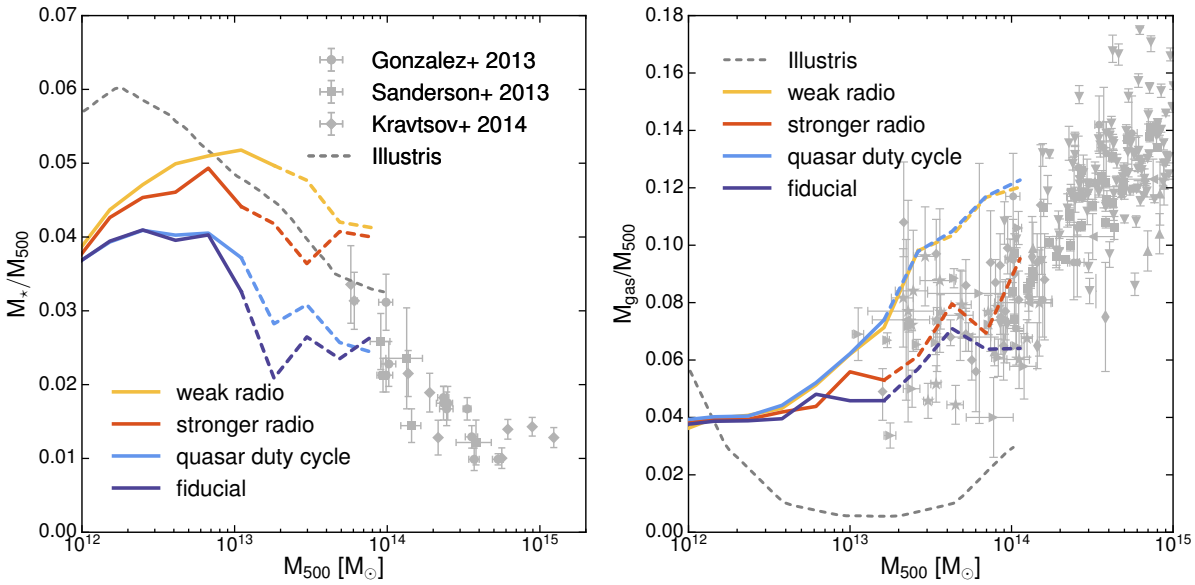


Figure A.2: Stellar mass (left) and gas mass (right) fractions inside r_{500} as a function of halo mass at $z = 0$ for different AGN feedback models. Lines show the mean relation in halo mass bins of width 0.2 dex. Lines become dashed when there are fewer than 10 haloes per bin. The grey dashed line shows the mean relation in Illustris. The total gas mass excludes cold and multiphase gas as described in Section 2.3.2. The observational data (symbols with error bars) are as shown in Fig. 2.4 and 2.5.

is, stronger radio-mode feedback slightly reduces the total stellar mass of massive haloes but the introduction of a quasar-mode duty cycle has a significantly larger effect.

In the right hand panel of Fig. A.2 we compare the median gas mass fraction as a function of halo mass for each model at $z = 0$. The 'stronger radio' model yields considerably lower gas fractions than the 'weak radio' model. This is owed to its more energetic bubble injections, which are able to eject gas from massive haloes more efficiently. This also explains why all four models yield much larger gas fractions than the Illustris model, which injected approximately an order of magnitude more energy per bubble feedback event. The 'quasar duty cycle' model yields gas mass fractions in massive haloes almost identical to the 'weak radio' model, however, because the 'weak radio' run converts significantly more gas into stars, the total mass of baryons accumulated by $z = 0$ is lower in the model with a quasar-mode duty cycle. This is because a quasar-mode duty cycle slows the accumulation of gas onto the virialised region of the halo by heating the ICM to higher temperatures. This effect is consistent with SPH simulations such as Le Brun et al. (2014), which show that discontinuous thermal AGN feedback can have a strong impact on the gas content of massive haloes.

In summary, we find that periodic quasar-mode AGN feedback significantly reduces stellar mass buildup in massive haloes compared to the continuous case. This has allowed us to reproduce the observed abundance of galaxies at the high mass end of the present-day GSMF without resorting to extremely strong radio-mode feedback, which was responsible for ejecting too much gas from massive haloes in the Illustris model. Using a far gentler form of radio-mode feedback in combination with a quasar-mode duty cycle allows us to reproduce the $z = 0$ GSMF and the local gas fractions of massive haloes simultaneously.

BIBLIOGRAPHY

Abazajian K. N., et al., 2016, preprint (arXiv:1610.02743)

Abell G. O., 1958, ApJS, 3, 211

Abell G. O., Corwin, Harold G. J., Olowin R. P., 1989, ApJS, 70, 1

Agertz O., Kravtsov A. V., 2016, ApJ, 824, 79

Agertz O., et al., 2007, MNRAS, 380, 963

Aihara H., et al., 2018, Publ. Astron. Soc. Japan, 70, S4

Aird J., et al., 2010, MNRAS, 401, 2531

Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, MNRAS, 451, 1892

Aird J., Coil A. L., Georgakakis A., 2018, MNRAS, 474, 1225

Akritas M. G., Bershady M. A., 1996, ApJ, 470, 706

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

Allgood B., Flores R. A., Primack J. R., Kravtsov A. V., Wechsler R. H., Faltenbacher A., Bullock J. S., 2006, MNRAS, 367, 1781

Anderson M. E., Gaspari M., White S. D. M., Wang W., Dai X., 2015, MNRAS, 449, 3806

Andersson K., et al., 2011, ApJ, 738, 48

Andrade-Santos F., et al., 2017, ApJ, 843, 76

Andreon S., Trinchieri G., Moretti A., Wang J., 2017, A&A, 606, A25

Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046

Applegate D. E., et al., 2014, MNRAS, 439, 48

Applegate D. E., et al., 2016, MNRAS, 457, 1522

Arnaud K. A., 1996, in Jacoby G. H., Barnes J., eds, Astronomical Society of the Pacific Conference Series Vol. 101, Astron. Data Anal. Softw. Syst. V. p. 17

Arnaud M., Pointecouteau E., Pratt G. W., 2007, A&A, 474, L37

Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92

Ascasibar Y., Yepes G., Gottlöber S., Müller V., 2004, MNRAS, 352, 1109

Ascaso B., Aguerri J. A. L., Varela J., Cava A., Bettoni D., Moles M., D'Onofrio M., 2011, ApJ, 726, 69
Ashby M. L. N., et al., 2013, ApJS, 209, 22
Babyk I. V., McNamara B. R., Nulsen P. E. J., Hogan M. T., Vantyghem A. N., Russell H. R., Pulido F. A., Edge A. C., 2018, ApJ, 857, 32
Bahé Y. M., McCarthy I. G., King L. J., 2012, MNRAS, 421, 1073
Bahé Y. M., et al., 2017, MNRAS, 470, 4186
Bai L., et al., 2014, ApJ, 789, 134
Baldry I. K., et al., 2012, MNRAS, 421, 621
Banerjee N., Sharma P., 2014, MNRAS, 443, 687
Barcons X., et al., 2017, Astron. Nachrichten, 338, 153
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Barnes J., Hut P., 1986, Nature, 324, 446
Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017a, MNRAS, 465, 213
Barnes D. J., et al., 2017b, MNRAS, 471, 1088
Barnes D. J., et al., 2018a, preprint (arXiv:1805.04109)
Barnes D. J., et al., 2018b, MNRAS, 481, 1809
Baron D., Netzer H., Poznanski D., Prochaska J. X., Förster Schreiber N. M., 2017, MNRAS, 470, 1687
Barret D., et al., 2018, in den Herder J.-W. A., Nakazawa K., Nikzad S., eds, Sp. Telesc. Instrum. 2018 Ultrav. to Gamma Ray. SPIE, p. 106991G (arXiv:1807.06092)
Bartalucci I., Arnaud M., Pratt G. W., Démoclès J., van der Burg R. F. J., Mazzotta P., 2017, A&A, 598, A61
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2012, ApJ, 758, 74
Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2013, ApJ, 777, 123
Battaglia N., et al., 2016, J. Cosmol. Astropart. Phys., 2016, 013
Baugh C. M., 2006, Reports Prog. Phys., 69, 3101
Baugh C. M., Gaztanaga E., Efstathiou G., 1995, MNRAS, 274, 1049
Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25
Bellstedt S., et al., 2016, MNRAS, 460, 2862
Bender A. N., et al., 2018, in Zmuidzinas J., Gao J.-R., eds, Millimeter, Submillimeter, Far-Infrared Detect. Instrum. Astron. IX. SPIE, p. 1070803 (arXiv:1809.00036)
Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001, MNRAS, 320, 261

Benson B. A., et al., 2014, in Holland W. S., Zmuidzinas J., eds, Millimeter, Submillimeter, Far-Infrared Detect. Instrum. Astron. VII. p. 91531P (arXiv:1407.2973)
Bernardi M., 2009, MNRAS, 395, 1491
Bernardi M., Hyde J. B., Sheth R. K., Miller C. J., Nichol R. C., 2007, Astron. J., 133, 1741
Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, MNRAS, 404, 2087
Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697
Bernardi M., Meert A., Sheth R. K., Fischer J. L., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2017, MNRAS, 467, 2217
Bertschinger E., 1985, ApJS, 58, 39
Bharadwaj V., Reiprich T. H., Lovisari L., Eckmiller H. J., 2015, A&A, 573, A75
Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122
Biffi V., Sembolini F., De Petris M., Valdarnini R., Yepes G., Gottlöber S., 2014, MNRAS, 439, 588
Biffi V., et al., 2016, ApJ, 827, 112
Biffi V., et al., 2017, MNRAS, 468, 531
Binney J., 1977, ApJ, 215, 483
Binney J., Tabor G., 1995, MNRAS, 276, 663
Birkinshaw M., 1999, Phys. Rep., 310, 97
Blanton M. R., et al., 2001, Astron. J., 121, 2358
Bleem L. E., et al., 2015, ApJS, 216, 27
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27
Bocquet S., et al., 2015, ApJ, 799, 214
Bocquet S., et al., 2018, preprint (arXiv:1812.01679)
Böhringer H., Voges W., Fabian A. C., Edge A. C., Neumann D. M., 1993, MNRAS, 264, L25
Böhringer H., et al., 2007, A&A, 469, 363
Bonafede A., Dolag K., Stasyszyn F., Murante G., Borgani S., 2011, MNRAS, 418, 2234
Bondi H., Hoyle F., 1944, MNRAS, 104, 273
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Borgani S., Kravtsov A., 2011, Adv. Sci. Lett., 4, 204
Bourne M. A., Sijacki D., 2017, MNRAS, 472, 4707

Bourne M. A., Sijacki D., Puchwein E., 2019, preprint (arXiv:1901.11030v1)
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Bower R. G., McCarthy I. G., Benson A. J., 2008, MNRAS, 390, 1399
Boyle Terlevich R. J., 1998, MNRAS, 293, L49
Brown M. J. I., et al., 2008, ApJ, 682, 937
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Budzynski J. M., Koposov S. E., McCarthy I. G., Belokurov V., 2013, MNRAS, 437, 1362
Bulbul E., et al., 2019, ApJ, 871, 50
Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517
Burenin R. A., Vikhlinin A., Hornstrup A., Ebeling H., Quintana H., Mescheryakov A., 2007, ApJS, 172, 561
Burke C., Collins C. A., 2013, MNRAS, 434, 2856
Burke C., Hilton M., Collins C., 2015, MNRAS, 449, 2353
Cano-Díaz M., Maiolino R., Marconi A., Netzer H., Shemmer O., Cresci G., 2012, A&A, 537, L8
Cappellari M., et al., 2006, MNRAS, 366, 1126
Carlstrom J. E., Joy M. K., Grego L., Holder G. P., Holzapfel W. L., Mohr J. J., Patel S., Reese E. D., 2000, Phys. Scr., T85, 148
Carlstrom J. E., Holder G. P., Reese E. D., 2002, ARA&A, 40, 643
Chabrier G., 2003, Publ. Astron. Soc. Pacific, 115, 763
Chisari N. E., et al., 2018, MNRAS, 480, 3962
Chiu I., et al., 2016a, MNRAS, 455, 258
Chiu I., et al., 2016b, MNRAS, 458, 379
Chiu I., et al., 2018, MNRAS, 478, 3072
Christensen C. R., Davé R., Governato F., Pontzen A., Brooks A., Munshi F., Quinn T., Wadsley J., 2016, ApJ, 824, 57
Churazov E., Bruggen M., Kaiser C. R., Böhringer H., Forman W., 2001, ApJ, 554, 261
Churazov E., Sazonov S., Sunyaev R., Forman W., Jones C., Böhringer H., 2005, MNRAS, 363, L91
Clerc N., et al., 2014, MNRAS, 444, 2723
Cohn J. D., White M., 2008, MNRAS, 385, 2025
Cole S., 1991, ApJ, 367, 45
Cole S., Lacey C., 1996, MNRAS, 281, 716
Collins C. A., et al., 2009, Nature, 458, 603

Conroy C., 2013, ARA&A, 51, 393
Conroy C., Ostriker J. P., 2008, ApJ, 681, 151
Conroy C., Wechsler R. H., Kravtsov A. V., 2007, ApJ, 668, 826
Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486
Contini E., De Lucia G., Villalobos A., Borgani S., 2014, MNRAS, 437, 3787
Cooke R. J., Pettini M., Jorgenson R. A., Murphy M. T., Steidel C. C., 2014, ApJ, 781, 31
Cooke K. C., O'Dea C. P., Baum S. A., Tremblay G. R., Cox I. G., Gladders M., 2016, ApJ, 833, 224
Cooke K. C., Fogarty K., Kartaltepe J. S., Moustakas J., O'Dea C. P., Postman M., 2018, ApJ, 857, 122
Cooper A. P., Gao L., Guo Q., Frenk C. S., Jenkins A., Springel V., White S. D. M., 2015, MNRAS, 451, 2703
Corless V. L., King L. J., 2007, MNRAS, 380, 149
Costa T., Rosdahl J., Sijacki D., Haehnelt M. G., 2018a, MNRAS, 473, 4197
Costa T., Rosdahl J., Sijacki D., Haehnelt M. G., 2018b, MNRAS, 479, 2079
Courtin J., Rasera Y., Alimi J.-M., Corasaniti P.-S., Boucher V., Füzfa A., 2010, MNRAS, 410, 1911
Crain R. A., et al., 2009, MNRAS, 399, 1773
Crain R. A., et al., 2015, MNRAS, 450, 1937
Crocce M., Fosalba P., Castander F. J., Gaztañaga E., 2010, MNRAS, 403, 1353
Croston J. H., et al., 2008, A&A, 487, 431
Croton D. J., et al., 2006, MNRAS, 365, 11
Cui W., et al., 2014a, MNRAS, 437, 816
Cui W., Borgani S., Murante G., 2014b, MNRAS, 441, 1769
Cui W., et al., 2018, MNRAS, 480, 2898
Cusworth S. J., Kay S. T., Battye R. A., Thomas P. A., 2014, MNRAS, 439, 2485
D'Souza R., Vegetti S., Kauffmann G., 2015, MNRAS, 454, 4027
Dark Energy Survey Collaboration 2005, preprint (arXiv:0510346)
Dark Energy Survey Collaboration 2016, MNRAS, 460, 1270
Davé R., Oppenheimer B. D., Sivanandam S., 2008, MNRAS, 391, 110
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Davison A. C., Hinkley D. V., 1997, Bootstrap methods and their application. Cambridge University Press, Cambridge
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
DeMaio T., Gonzalez A. H., Zabludoff A., Zaritsky D., Connor T., Donahue M., Mulchaey J. S., 2018, MNRAS, 474, 3009
Decker B., et al., 2019, preprint (arXiv:1903.11100)

Del Popolo A., 2012, MNRAS, 424, 38
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Dietrich J. P., et al., 2019, MNRAS, 483, 2871
Dijkstra M., Haiman Z., Rees M. J., Weinberg D. H., 2004, ApJ, 601, 666
Dolag K., Vazza F., Brunetti G., Tormen G., 2005, MNRAS, 364, 753
Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
Dolag K., Murante G., Borgani S., 2010, MNRAS, 405, 1544
Dolag K., Komatsu E., Sunyaev R., 2016, MNRAS, 463, 1797
Donahue M., et al., 2014, ApJ, 794, 136
Donahue M., et al., 2015, ApJ, 805, 177
Dubois Y., Devriendt J., Teyssier R., Slyz A., 2011, MNRAS, 417, 1853
Dubois Y., et al., 2014, MNRAS, 444, 1453
Dunn R. J. H., Fabian A. C., 2006, MNRAS, 373, 959
Dunn R. J. H., Fabian A. C., 2008, MNRAS, 385, 757
Dunn R. J. H., Fabian A. C., Taylor G. B., 2005, MNRAS, 364, 1343
Eckert D., et al., 2016, A&A, 592, A12
Eckmiller H. J., Hudson D. S., Reiprich T. H., 2011, A&A, 535, A105
Edgar R., 2004, New Astron. Rev., 48, 843
Edge A., 2001, MNRAS, 328, 762
Edge A. C., Frayer D. T., 2003, ApJ, 594, L13
Efstathiou G., 1992, MNRAS, 256, 43P
Efstathiou G., Silk J., 1983, Fundam. Cosm. Phys., 9, 1
Efstathiou G., Davis M., White S. D. M., Frenk C. S., 1985, ApJS, 57, 241
Einasto J., Saar E., Kaasik A., Chernin A. D., 1974, Nature, 252, 111
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Eisenstein D. J., Hu W., 1999, ApJ, 511, 5
English W., Hardcastle M. J., Krause M. G. H., 2016, MNRAS, 461, 2025
Erben T., et al., 2013, MNRAS, 433, 2545
Ettori S., Tozzi P., Borgani S., Rosati P., 2004, A&A, 417, 13
Fabian A. C., 2012, ARA&A, 50, 455
Fabian A. C., Sanders J. S., Allen S. W., Crawford C. S., Iwasawa K., Johnstone R. M., Schmidt R. W., Taylor G. B., 2003, MNRAS, 344, L43
Fabian A. C., Walker S. A., Russell H. R., Pinto C., Sanders J. S., Reynolds C. S., 2017, MNRAS, 464, L1
Fabjan D., Borgani S., Tornatore L., Saro A., Murante G., Dolag K., 2010, MNRAS, 401, 1670
Fabjan D., Borgani S., Rasia E., Bonafede A., Dolag K., Murante G., Tornatore L., 2011, MNRAS, 416, 801

Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, MNRAS, 406, 2267
Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2018, MNRAS, 478, 2618
Farrah D., et al., 2012, ApJ, 745, 178
Fillmore J. A., Goldreich P., 1984, ApJ, 281, 1
Foëx G., Böhringer H., Chon G., 2017, A&A, 606, A122
Fogarty K., Postman M., Connor T., Donahue M., Moustakas J., 2015, ApJ, 813, 117
Fraser-McKelvie A., Brown M. J. I., Pimbblet K. A., 2014, MNRAS, 444, L63
Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, ApJ, 327, 507
Frenk C. S., et al., 1999, ApJ, 525, 554
Frieman J. A., Turner M. S., Huterer D., 2008, ARA&A, 46, 385
Fujita Y., Suzuki T. K., Wada K., 2004, ApJ, 600, 650
Furlong M., et al., 2017, MNRAS, 465, 722
Furnell K. E., et al., 2018, MNRAS, 478, 4952
Gardner J. P., et al., 2006, Space Sci. Rev., 123, 485
Gaspari M., Brighenti F., Temi P., Ettori S., 2014, ApJ, 783, L10
Gastaldello F., Buote D. A., Humphrey P. J., Zappacosta L., Bullock J. S., Brighenti F., Mathews W. G., 2007, ApJ, 669, 158
Ge C., Sun M., Rozo E., Sehgal N., Vikhlinin A., Forman W., Jones C., Nagai D., 2019, MNRAS, 484, 1946
Geller M. J., Huchra J. P., 1989, Science, 246, 897
Genel S., et al., 2014, MNRAS, 445, 175
George M. R., et al., 2012, ApJ, 757, 2
Giles P. A., et al., 2016, A&A, 592, A3
Giles P. A., et al., 2017, MNRAS, 465, 858
Giodini S., et al., 2009, ApJ, 703, 982
Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16
Gonzalez A. H., Zabludoff A. I., Zaritsky D., 2005, ApJ, 618, 195
Gonzalez A. H., Zaritsky D., Zabludoff A. I., 2007, ApJ, 666, 147
Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, ApJ, 778, 14
Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Gozaliasl G., Finoguenov A., Khosroshahi H. G., Mirkazemi M., Erfanianfar G., Tanaka M., 2016, MNRAS, 458, 2762
Gozaliasl G., et al., 2018, MNRAS, 475, 2787
Graham A., Lauer T. R., Colless M., Postman M., 1996, ApJ, 465, 534
Graham A. W., Driver S. P., Petrosian V., Conselice C. J., Bershady M. A., Crawford S. M., Goto T., 2005, Astron. J., 130, 1535

Graham A. W., Merritt D., Moore B., Diemand J., Terzić B., 2006, Astron. J., 132, 2701
Grand R. J. J., et al., 2017, MNRAS, 467, 179
Greco J. P., Colin Hill J., Spergel D. N., Battaglia N., 2015, ApJ, 808, 151
Greene J. E., Zakamska N. L., Ho L. C., Barth A. J., 2011, ApJ, 732, 9
Greene J. E., Pooley D., Zakamska N. L., Comerford J. M., Sun A.-L., 2014, ApJ, 788, 54
Gruen D., et al., 2014, MNRAS, 442, 1507
Guedes J., Callegari S., Madau P., Mayer L., 2011, ApJ, 742, 76
Gunn J. E., Gott J., 1972, ApJ, 176, 1
Guo F., Oh S. P., 2008, MNRAS, 384, 251
Guo Q., et al., 2011, MNRAS, 413, 101
Guo Q., et al., 2016, MNRAS, 461, 3457
Guth A. H., 1981, Phys. Rev. D, 23, 347
Hahn O., Martizzi D., Wu H.-Y. Y., Evrard A. E., Teyssier R., Wechsler R. H., 2017, MNRAS, 470, 166
Haring N., Rix H.-W., 2004, ApJ, 604, L89
Harrison C. M., Alexander D. M., Mullaney J. R., Swinbank A. M., 2014, MNRAS, 441, 3306
Heath D. J., 1977, MNRAS, 179, 351
Heitmann K., et al., 2008, Comput. Sci. Discov., 1, 015003
Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., Pearce F. R., 2003, MNRAS, 338, 913
Helsdon S. F., Ponman T. J., 2000, MNRAS, 315, 356
Henderson S. W., et al., 2016, J. Low Temp. Phys., 184, 772
Henson M. A., Barnes D. J., Kay S. T., McCarthy I. G., Schaye J., 2017, MNRAS, 465, 3361
Heymans C., et al., 2012, MNRAS, 427, 146
Hillel S., Soker N., 2017a, MNRAS, 466, L39
Hillel S., Soker N., 2017b, ApJ, 845, 91
Hilton M., et al., 2012, MNRAS, 424, 2086
Hirschmann M., Naab T., Somerville R. S., Burkert A., Oser L., 2012, MNRAS, 419, 3200
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles. McGraw-Hill, New York
Hoekstra H., Hartlap J., Hilbert S., van Uitert E., 2011, MNRAS, 412, 2095

Hoekstra H., Bartelmann M., Dahle H. k., Israel H., Limousin M., Meneghetti M., 2013, Space Sci. Rev., 177, 75
Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS, 449, 685
Hoffer A. S., Donahue M., Hicks A., Barthelemy R. S., 2012, ApJS, 199, 23
Hoffman Y., Shaham J., 1985, ApJ, 297, 16
Holtzman J. A., 1989, ApJS, 71, 1
Hopkins P. F., Bundy K., Murray N., Quataert E., Lauer T. R., Ma C.-P., 2009, MNRAS, 398, 898
Hopkins P. F., Bundy K., Hernquist L., Wuyts S., Cox T. J., 2010, MNRAS, 401, 1099
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Hoyle F., Lyttleton R. A., 1939, Proc. Camb. Philol. Soc., 35, 405
Huang S., et al., 2018, preprint (arXiv:1811.01139)
Hubble E. P., 1926, ApJ, 64, 321
Hudson D. S., Mittal R., Reiprich T. H., Nulsen P. E. J., Andernach H., Sarazin C. L., 2010, A&A, 513, A37
Hurier G., Angulo R. E., 2018, A&A, 610, L4
Ilbert O., et al., 2013, A&A, 556, A55
Ishibashi W., Fabian A. C., Maiolino R., 2018, MNRAS, 476, 512
Ishiyama T., Enoki M., Kobayashi M. A. R., Makiya R., Nagashima M., Oogi T., 2015, Publ. Astron. Soc. Japan, 67, 61
Isobe T., Feigelson E. D., Akritas M. G., Babu G. J., 1990, ApJ, 364, 104
Israel H., Reiprich T. H., Erben T., Massey R. J., Sarazin C. L., Schneider P., Vikhlinin A., 2014, A&A, 564, A129
Israel H., Schellenberger G., Nevalainen J., Massey R., Reiprich T. H., 2015, MNRAS, 448, 814
Jacob S., Pfrommer C., 2017, MNRAS, 467, 1478
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Jiménez-Teja Y., et al., 2018, ApJ, 857, 79
Jing Y. P., Suto Y., 2002, ApJ, 574, 538
Johnstone R. M., Fabian A. C., Nulsen P. E. J., 1987, MNRAS, 224, 75
Kaiser N., 1986, MNRAS, 222, 323
Kannan R., Stinson G. S., Macciò A. V., Brook C., Weinmann S. M., Wadsley J., Couchman H. M. P., 2014, MNRAS, 437, 3529
Kannan R., Springel V., Pakmor R., Marinacci F., Vogelsberger M., 2016, MNRAS, 458, 410

Kannan R., Vogelsberger M., Pfrommer C., Weinberger R., Springel V., Hernquist L., Puchwein E., Pakmor R., 2017, ApJ, 837, L18
Katz N., Gunn J. E., 1991, ApJ, 377, 365
Katz N., White S. D. M., 1993, ApJ, 412, 455
Katz N., Hernquist L., Weinberg D. H., 1992, ApJ, 399, L109
Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Kay S. T., Thomas P. A., Jenkins A., Pearce F. R., 2004, MNRAS, 355, 1091
Kelly B. C., 2007, ApJ, 665, 1489
Kelly P. L., et al., 2014, MNRAS, 439, 28
Kereš D., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, MNRAS, 425, 2027
Kettula K., et al., 2013, ApJ, 778, 74
Kettula K., et al., 2015, MNRAS, 451, 1460
Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M. S., 2015, MNRAS, 450, 1349
Khedekar S., Churazov E., Kravtsov A., Zhuravleva I., Lau E. T., Nagai D., Sunyaev R., 2013, MNRAS, 431, 954
King A., Pounds K., 2015, ARA&A, 53, 115
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kotov O., Vikhlinin A., 2005, ApJ, 633, 781
Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353
Kravtsov A. V., Vikhlinin A., Nagai D., 2006, ApJ, 650, 128
Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astron. Lett., 44, 8
Krick J. E., Bernstein R. A., 2007, Astron. J., 134, 466
Kroupa P., 2001, MNRAS, 322, 231
LSST Dark Energy Science Collaboration 2012, preprint (arXiv:1211.0310)
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lacy M., et al., 2019, MNRAS, 483, L22
Lansbury G. B., Jarvis M. E., Harrison C. M., Alexander D. M., Moro A. D., Edge A. C., Mullaney J. R., Thomson A. P., 2018, ApJ, 856, L1
Lau E. T., Kravtsov A. V., Nagai D., 2009, ApJ, 705, 1129
Lau E. T., Nagai D., Nelson K., 2013, ApJ, 777, 151
Laureijs R., et al., 2011, preprint (arXiv:1110.3193)
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270
Le Brun A. M., McCarthy I. G., Melin J. B., 2015, MNRAS, 451, 3868
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2016, MNRAS, 466, 4442
Leauthaud A., et al., 2010, ApJ, 709, 97

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Li C., White S. D. M., 2009, MNRAS, 398, 2177
Li Y., Ruszkowski M., Bryan G. L., 2017, ApJ, 847, 106
Lidman C., et al., 2012, MNRAS, 427, 550
Lieu M., et al., 2016, A&A, 592, A4
Lim S. H., Mo H. J., Li R., Liu Y., Ma Y.-Z., Wang H., Yang X., 2018, ApJ, 854, 181
Lin Y.-T., Mohr J. J., 2004, ApJ, 617, 879
Lin C. C., Mestel L., Shu F. H., 1965, ApJ, 142, 1431
Lin Y.-T., Mohr J. J., Stanford S. A., 2003, ApJ, 591, 749
Lin Y.-T., Stanford S. A., Eisenhardt P., Vikhlinin A., Maughan B. J., Kravtsov A., 2012, Astrophys. J. Lett., 745, 6
Lin Y.-T., Brodwin M., Gonzalez A. H., Bode P., Eisenhardt P. R., Stanford S. A., Vikhlinin A., 2013, ApJ, 771
Lin Y.-T., et al., 2017, ApJ, 851, 139
Liu G., Zakamska N. L., Greene J. E., Nesvadba N. P. H., Liu X., 2013, MNRAS, 436, 2576
Louis T., Alonso D., 2017, Phys. Rev. D, 95, 043517
Lovisari L., Reiprich T. H., Schellenberger G., 2015, A&A, 573, A118
Lovisari L., et al., 2017, ApJ, 846, 51
Lynden-Bell D., 1967, MNRAS, 136, 101
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Mahdavi A., Hoekstra H., Babul A., Henry J. P., 2008, MNRAS, 384, 1567
Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, ApJ, 767, 116
Mancone C. L., Gonzalez A. H., 2012, Publ. Astron. Soc. Pacific, 124, 606
Mandelbaum R., 2018, ARA&A, 56, 393
Manrique A., Raig A., Salvador-Sole E., Sanchis T., Solanes J. M., 2003, ApJ, 593, 26
Mantz A., Allen S. W., Ebeling H., Rapetti D., Drlica-Wagner A., 2010, MNRAS, 406, 1773
Mantz A. B., et al., 2015, MNRAS, 446, 2205
Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., 2016a, MNRAS, 456, 4020
Mantz A. B., et al., 2016b, MNRAS, 463, 3582
Marian L., Smith R. E., Bernstein G. M., 2010, ApJ, 709, 286
Marinacci F., Pakmor R., Springel V., 2014, MNRAS, 437, 1750
Marinacci F., et al., 2018, MNRAS, 480, 5113
Markevitch M., 1998, ApJ, 504, 27
Martizzi D., Teyssier R., Moore B., 2012, MNRAS, 420, 2859
Martizzi D., Jimmy Teyssier R., Moore B., 2014, MNRAS, 443, 1500

Mathiesen B. F., Evrard A. E., 2001, ApJ, 546, 100
Maughan B. J., 2007, ApJ, 668, 772
Maughan B. J., 2013, MNRAS, 437, 1171
Maughan B. J., Jones L. R., Ebeling H., Scharf C., 2006, MNRAS, 365, 509
Maughan B. J., Jones C., Forman W., Van Speybroeck L., 2008, ApJS, 174, 117
Maughan B. J., Giles P. A., Randall S. W., Jones C., Forman W. R., 2012, MNRAS, 421, 1583
Maughan B. J., Giles P. A., Rines K. J., Diaferio A., Geller M. J., Van Der Pyl N., Bonamente M., 2016, MNRAS, 461, 4182
Mazzotta P., Rasia E., Moscardini L., Tormen G., 2004, MNRAS, 354, 10
McCarthy I. G., et al., 2010, MNRAS, 406, 822
McCarthy I. G., Schaye J., Bower R. G., Ponman T. J., Booth C. M., Vecchia C. D., Springel V., 2011, MNRAS, 412, 1965
McCarthy I. G., Schaye J., Font A. S., Theuns T., Frenk C. S., Crain R. A., Dalla Vecchia C., 2012, MNRAS, 427, 379
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
McDonald M., et al., 2016, ApJ, 817, 86
McDonald M., Gaspari M., McNamara B. R., Tremblay G. R., 2018, ApJ, 858, 45
McGee S. L., Balogh M. L., 2010, MNRAS, 403, L79
McNamara B., Nulsen P., 2007, ARA&A, 45, 117
McNamara B. R., Nulsen P. E. J., Wise M. W., Rafferty D. A., Carilli C., Sarazin C. L., Blanton E. L., 2005, Nature, 433, 45
Medezinski E., Broadhurst T., Umetsu K., Oguri M., Rephaeli Y., Benítez N., 2010, MNRAS, 405, 257
Medezinski E., et al., 2018, Publ. Astron. Soc. Japan, 70, S28
Mehrtens N., et al., 2012, MNRAS, 423, 1024
Melin J.-B., Bartlett J. G., 2015, A&A, 578, A21
Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, A&A, 514, A93
Merloni A., Predehl P., Becker W., Böhringer H., Boller T., Brunner H., Brusa M., 2012, preprint (arXiv:1209.3114)
Miralda-Escude J., Babul A., 1995, ApJ, 449, 18
Mitchell N. L., McCarthy I. G., Bower R. G., Theuns T., Crain R. A., 2009, MNRAS, 395, 180
Mitchell P. D., et al., 2018, MNRAS, 474, 492
Mittal R., et al., 2011, MNRAS, 418, 2386
Mittal R., Whelan J. T., Combes F., 2015, MNRAS, 450, 2564

Mo H., van den Bosch F., White S., 2010, Galaxy Formation and Evolution. Cambridge University Press, Cambridge
Montes M., Trujillo I., 2018, MNRAS, 474, 917
Morandi A., Limousin M., 2012, MNRAS, 421, 3147
Morishita T., Abramson L. E., Treu T., Schmidt K. B., Vulcani B., Wang X., 2017, ApJ, 846, 139
Mulchaey J. S., 2000, ARA&A, 38, 289
Murante G., Giovalli M., Gerhard O., Arnaboldi M., Borgani S., Dolag K., 2007, MNRAS, 377, 2
Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569
Murray S. G., Power C., Robotham A. S., 2013a, Astron. Comput., 3-4, 23
Murray S. G., Power C., Robotham A. S. G., 2013b, MNRAS, 434, L61
Mushotzky R. F., 2018, in den Herder J.-W. A., Nakazawa K., Nikzad S., eds, Sp. Telesc. Instrum. 2018 Ultrav. to Gamma Ray. SPIE, p. 80
Muzzin A., et al., 2013, ApJ, 777, 18
Nagai D., Vikhlinin A., Kravtsov A. V., 2007a, ApJ, 655, 98
Nagai D., Kravtsov A. V., Vikhlinin A., 2007b, ApJ, 668, 1
Nagarajan A., et al., 2018, MNRAS tmp, 1811
Naiman J. P., et al., 2018, MNRAS, 477, 1206
Nandra K., et al., 2013, preprint (arXiv:1306.2307)
Navarro J. F., White S. D. M., 1994, MNRAS, 267, 401
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Navarro J. F., et al., 2004, MNRAS, 349, 1039
Nelson D., Vogelsberger M., Genel S., Sijacki D., Kereš D., Springel V., Hernquist L., 2013, MNRAS, 429, 3353
Nelson K., Lau E. T., Nagai D., Rudd D. H., Yu L., 2014a, ApJ, 782, 107
Nelson K., Lau E. T., Nagai D., 2014b, ApJ, 792, 25
Nelson D., et al., 2018, MNRAS, 475, 624
Nelson D., et al., 2019, preprint (arXiv:1902.05554)
Newman A. B., Treu T., Ellis R. S., Sand D. J., 2011, ApJ, 728, L39
Nusser A., Sheth R. K., 1999, MNRAS, 303, 685
O'Dea C. P., et al., 2008, ApJ, 681, 1035
Oguri M., Takada M., Umetsu K., Broadhurst T., 2005, ApJ, 632, 841
Okabe N., Smith G. P., 2016, MNRAS, 461, 3794
Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920
Omma H., Binney J., Bryan G., Slyz A., 2004, MNRAS, 348, 1105
Osmond J. P. F., Ponman T. J., 2004, MNRAS, 350, 1511
Ostriker J. P., Peebles P. J. E., Yahil A., 1974, ApJ, 193, L1

Ota N., Nagai D., Lau E. T., 2018, Publ. Astron. Soc. Japan, 70, 51
O'Shea B. W., Nagamine K., Springel V., Hernquist L., Norman M. L., 2005, ApJS, 160, 1
Pacaud F., et al., 2007, MNRAS, 382, 1289
Pacaud F., et al., 2016, A&A, 592, A2
Pakmor R., Pfrommer C., Simpson C. M., Springel V., 2016, ApJ, 824, L30
Parroni C., et al., 2017, ApJ, 848, 114
Peebles P. J. E., 1980, The large-scale structure of the Universe. Princeton Univ. Press, Princeton, NJ
Penna-Lima M., Bartlett J. G., Rozo E., Melin J.-B., Merten J., Evrard A. E., Postman M., Rykoff E., 2017, A&A, 604, A89
Peterson J. R., et al., 2001, A&A, 365, L104
Peterson J. R., Kahn S. M., Paerels F. B. S., Kaastra J. S., Tamura T., Bleeker J. A. M., Ferrigno C., Jernigan J. G., 2003, ApJ, 590, 207
Pfrommer C., 2013, ApJ, 779, 10
Pfrommer C., Pakmor R., Schaal K., Simpson C. M., Springel V., 2017, MNRAS, 465, 4500
Piffaretti R., Valdarnini R., 2008, A&A, 491, 71
Pike S. R., Kay S. T., Newton R. D. A., Thomas P. A., Jenkins A., 2014, MNRAS, 445, 1774
Pillepich A., Porciani C., Reiprich T. H., 2012, MNRAS, 422, 44
Pillepich A., et al., 2018a, MNRAS, 473, 4077
Pillepich A., et al., 2018b, MNRAS, 475, 648
Pillepich A., et al., 2019, preprint (arXiv:1902.05553)
Planck Collaboration IX 2016, A&A, 594, A9
Planck Collaboration VI 2018, preprint (arXiv:1807.06209)
Planck Collaboration XI 2011, A&A, 536, A11
Planck Collaboration XI 2013, A&A, 557, A52
Planck Collaboration XIII 2016, A&A, 594, A13
Planck Collaboration XX 2014, A&A, 571, A20
Planck Collaboration XXIV 2016, A&A, 594, A24
Planck Collaboration XXIX 2014, A&A, 571, A29
Planck Collaboration XXVII 2016, A&A, 594, A27
Planelles S., Borgani S., Dolag K., Ettori S., Fabjan D., Murante G., Tornatore L., 2013, MNRAS, 431, 1487
Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195
Planelles S., Schleicher D. R. G., Bykov A. M., 2015, Space Sci. Rev., 188, 93

Planelles S., et al., 2017, MNRAS, 467, 3827
Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14
Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, A&A, 498, 361
Pratt G. W., et al., 2010, A&A, 511, 85
Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, Space Sci. Rev., 215, 25
Press W. H., Schechter P., 1974, ApJ, 187, 425
Price D. J., 2008, J. Comput. Phys., 227, 10040
Puchwein E., Springel V., 2013, MNRAS, 428, 2966
Puchwein E., Sijacki D., Springel V., 2008, ApJ, 687, L53
Puchwein E., Springel V., Sijacki D., Dolag K., 2010, MNRAS, 406, 936
Pulido F. A., et al., 2018, ApJ, 853, 177
Purcell C. W., Bullock J. S., Zentner A. R., 2007, ApJ, 666, 20
Quilis V., Bower R. G., Balogh M. L., 2001, MNRAS, 328, 1091
Ragone-Figueroa C., Granato G. L., Murante G., Borgani S., Cui W., 2013, MNRAS, 436, 1750
Ragone-Figueroa C., Granato G. L., Ferraro M. E., Murante G., Biffi V., Borgani S., Planelles S., Rasia E., 2018, MNRAS, 479, 1125
Randall S. W., et al., 2015, ApJ, 805, 112
Rasia E., Mazzotta P., Borgani S., Moscardini L., Dolag K., Tormen G., Diaferio A., Murante G., 2005, ApJ, 618, L1
Rasia E., et al., 2006, MNRAS, 369, 2013
Rasia E., et al., 2012, New J. Phys., 14, 055018
Rasia E., et al., 2014, ApJ, 791, 96
Rasia E., et al., 2015, ApJ, 813, L17
Rawle T. D., et al., 2012, ApJ, 747, 29
Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541
Reichert A., Böhringer H., Fassbender R., Mühlegger M., 2011, A&A, 535, A4
Reiprich T. H., 2017, Astron. Nachrichten, 338, 349
Remus R.-S., Dolag K., Hoffmann T., 2017, Galaxies, 5, 49
Richard J., et al., 2010, MNRAS, 404, 325
Riess A. G., et al., 2011, ApJ, 730, 119
Rines K. J., Geller M. J., Diaferio A., Hwang H. S., 2016, ApJ, 819, 63
Roberts M. S., Rots A. H., 1973, A&A, 26, 483
Roncarelli M., et al., 2018, A&A, 618, A39
Rosati P., Borgani S., Norman C., 2002, ARA&A, 40, 539
Rossetti M., et al., 2016, MNRAS, 457, 4515

Rossetti M., Gastaldello F., Eckert D., Della Torre M., Pantiri G., Cazzoletti P., Molendi S., 2017, MNRAS, 468, 1917
Rozo E., et al., 2009, ApJ, 699, 768
Rozo E., et al., 2010, ApJ, 708, 645
Rubin V. C., Thonnard N., Ford, W. K. J., 1978, ApJ, 225, L107
Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, ApJ, 672, 19
Rudick C. S., Mihos J. C., McBride C. K., 2011, ApJ, 732, 48
Rupke D. S. N., Gültekin K., Veilleux S., 2017, ApJ, 850, 40
Russell H. R., et al., 2014, ApJ, 784, 78
Russell H. R., et al., 2017, MNRAS, 472, 4024
Ruszkowski M., Begelman M. C., 2002, ApJ, 581, 223
Ruszkowski M., Oh S. P., 2010, ApJ, 713, 1332
Ruszkowski M., Oh S. P., 2011, MNRAS, 414, 1493
Ruszkowski M., Bruggen M., Begelman M. C., 2004, ApJ, 611, 158
Ruszkowski M., Lee D., Brüggen M., Parrish I., Oh S. P., 2011, ApJ, 740, 81
Ruszkowski M., Yang H.-Y. K., Reynolds C. S., 2017, ApJ, 844, 13
Rybicki G. B., Lightman A. P., 1979, Radiative Processes in Astrophysics. John Wiley & Sons, New York
Ryden B. S., Gunn J. E., 1987, ApJ, 318, 15
Saliwanchik B. R., et al., 2015, ApJ, 799, 137
Salomé P., Combes F., 2003, A&A, 412, 657
Salpeter E. E., 1955, ApJ, 121, 161
Sánchez C., et al., 2014, MNRAS, 445, 1482
Sanders J. S., Fabian A. C., Allen S. W., Morris R. G., Graham J., Johnstone R. M., 2008, MNRAS, 385, 1186
Sanderson A. J. R., Ponman T. J., Finoguenov A., Lloyd-Davies E. J., Markevitch M., 2003, MNRAS, 340, 989
Sanderson A. J. R., O'Sullivan E., Ponman T. J., Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, MNRAS, 429, 3288
Sarazin C. L., 1986, Rev. Mod. Phys., 58, 1
Sartori L. F., et al., 2016, MNRAS, 457, 3629
Schaye J., et al., 2010, MNRAS, 402, 1536
Schaye J., et al., 2015, MNRAS, 446, 521
Schellenberger G., Reiprich T. H., Lovisari L., Nevalainen J., David L., 2015, A&A, 575, A30
Schrabback T., et al., 2018, MNRAS, 474, 2635
Sehgal N., et al., 2011, ApJ, 732, 44
Seigar M. S., Graham A. W., Jerjen H., 2007, MNRAS, 378, 1575

Sembolini F., Yepes G., De Petris M., Gottlöber S., Lamagna L., Comis B., 2013, MNRAS, 429, 323
Sembolini F., De petris M., Yepes G., Foschi E., Lamagna L., Gottlöber S., 2014, MNRAS, 440, 3520
Sembolini F., et al., 2016a, MNRAS, 457, 4063
Sembolini F., et al., 2016b, MNRAS, 459, 2973
Semboloni E., Hoekstra H., Schaye J., Van Daalen M. P., Mccarthy I. G., 2011, MNRAS, 417, 2020
Sereno M., Ettori S., 2015a, MNRAS, 450, 3633
Sereno M., Ettori S., 2015b, MNRAS, 450, 3675
Sereno M., Giocoli C., Ettori S., Moscardini L., 2015a, MNRAS, 449, 2024
Sereno M., Ettori S., Moscardini L., 2015b, MNRAS, 450, 3649
Sereno M., Covone G., Izzo L., Ettori S., Coupon J., Lieu M., 2017, MNRAS, 472, 1946
Shen J., Abel T., Mo H. J., Sheth R. K., 2006, ApJ, 645, 783
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Shi X., Komatsu E., Nelson K., Nagai D., 2015, MNRAS, 448, 1020
Shi X., Komatsu E., Nagai D., Lau E. T., 2016, MNRAS, 455, 2936
Short C. J., Thomas P. A., Young O. E., Pearce F. R., Jenkins A., Muanwong O., 2010, MNRAS, 408, 2213
Sijacki D., Springel V., 2006, MNRAS, 366, 397
Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Sijacki D., Pfrommer C., Springel V., Enlin T. A., 2008, MNRAS, 387, 1403
Sijacki D., Vogelsberger M., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 424, 2999
Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575
Silverman J. D., et al., 2008, ApJ, 679, 118
Simet M., Battaglia N., Mandelbaum R., Seljak U., 2017, MNRAS, 466, 3663
Simpson C. M., Pakmor R., Marinacci F., Pfrommer C., Springel V., Glover S. C. O., Clark P. C., Smith R. J., 2016, ApJ, 827, L29
Smith S., 1936, ApJ, 83, 23
Smith R. K., Brickhouse N. S., Liedahl D. A., Raymond J. C., 2001, ApJ, 556, L91
Smith G. P., et al., 2016, MNRAS, 456, L74
Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481
Sommer-Larsen J., Romeo A. D., Portinari L., 2005, MNRAS, 357, 478

Sonnenfeld A., Treu T., Gavazzi R., Suyu S. H., Marshall P. J., Auger M. W., Nipoti C., 2013, ApJ, 777, 98

Spergel D., et al., 2015, preprint (arXiv:1503.03757)

Springel V., 2010a, ARA&A, 48, 391

Springel V., 2010b, MNRAS, 401, 791

Springel V., Hernquist L., 2003a, MNRAS, 339, 289

Springel V., Hernquist L., 2003b, MNRAS, 339, 312

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Springel V., Di Matteo T., Hernquist L., 2005a, MNRAS, 361, 776

Springel V., et al., 2005b, Nature, 435, 629

Springel V., et al., 2018, MNRAS, 475, 676

Stanek R., Rasia E., Evrard A. E., Pearce F., Gazzola L., 2010, ApJ, 715, 1508

Steinborn L. K., Dolag K., Hirschmann M., Prieto M. A., Remus R.-S., 2015, MNRAS, 448, 1504

Stott J. P., et al., 2010, ApJ, 718, 23

Stott J. P., Collins C. A., Burke C., Hamilton-Morris V., Smith G. P., 2011, MNRAS, 414, 445

Sun M., 2012, New J. Phys., 14, 045004

Sun M., Voit G. M., Donahue M., Jones C., Forman W., Vikhlinin A., 2009, ApJ, 693, 1142

Sunyaev R. A., Zel'dovich Y. B., 1970, Ap&SS, 7, 3

Sunyaev R. A., Zel'dovich Y. B., 1972, Comments Astrophys. Sp. Phys., 4, 173

Takahashi T., et al., 2010, Proc.SPIE, 7732, 77320Z

Takahashi T., et al., 2016, Proc.SPIE, 9905, 99050U

Tamura T., et al., 2001, A&A, 365, L87

Tamura T., Kaastra J. S., Makishima K., Takahashi I., 2003, A&A, 399, 497

Tashiro M., et al., 2018, in den Herder J.-W. A., Nakazawa K., Nikzad S., eds, Sp. Telesc. Instrum. 2018 Ultrav. to Gamma Ray. SPIE, p. 73

Tasker E. J., Brunino R., Mitchell N. L., Michielsen D., Hopton S., Pearce F. R., Bryan G. L., Theuns T., 2008, MNRAS, 390, 1267

The Lynx Team 2018, preprint (arXiv:1809.09642)

The Simons Observatory Collaboration 2019, J. Cosmol. Astropart. Phys., 2019, 056

Thoul A. A., Weinberg D. H., 1995, ApJ, 442, 480

Tinker J. L., et al., 2008, ApJ, 688, 709

Tomczak A. R., et al., 2014, ApJ, 783, 85

Tonnesen S., Bryan G. L., 2012, MNRAS, 422, 1609

Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, MNRAS, 438, 1985
Tremaine S., et al., 2002, ApJ, 574, 740
Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, MNRAS, 470, 1121
Truong N., et al., 2018, MNRAS, 474, 4089
Tyson J. A., Valdes F., Jarvis J. F., Mills, A. P. J., 1984, ApJ, 281, L59
Vanderlinde K., et al., 2010, ApJ, 722, 1180
Veilleux S., Meléndez M., Sturm E., Gracia-Carpio J., Fischer J., 2013, ApJ, 776, 27
Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Vecchia C. D., 2014, MNRAS, 442, 2641
Vikhlinin A., VanSpeybroeck L., Markevitch M., Forman W. R., Grego L., 2002, ApJ, 578, L107
Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691
Vikhlinin A., et al., 2009, ApJ, 692, 1033
Vitvitska M., Klypin A. A., Kravtsov A. V., Wechsler R. H., Primack J. R., Bullock J. S., 2002, ApJ, 581, 799
Vogelsberger M., Sijacki D., Kereš D., Springel V., Hernquist L., Keres D., Springel V., Hernquist L., 2012, MNRAS, 425, 3024
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Vogelsberger M., et al., 2014a, MNRAS, 444, 1518
Vogelsberger M., et al., 2014b, Nature, 509, 177
Vogelsberger M., et al., 2018, MNRAS, 474, 2073
Voigt L. M., Fabian A. C., 2004, MNRAS, 347, 1130
Voit G., 2005, Rev. Mod. Phys., 77, 207
Voit G. M., Kay S. T., Bryan G. L., 2005, MNRAS, 364, 909
Von Der Linden A., Best P. N., Kauffmann G., White S. D. M., 2007, MNRAS, 379, 867
Wadsley J. W., Veeravalli G., Couchman H. M. P., 2008, MNRAS, 387, 427
Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, MNRAS, 454, 83
Wang W., White S. D. M., Mandelbaum R., Henriques B., Anderson M. E., Han J., 2016, MNRAS, 456, 2301
Wang L., Xu D., Gao L., Guo Q., Qu Y., Pan J., 2019, MNRAS, 485, 2083
Warren M. S., Quinn P. J., Salmon J. K., Zurek W. H., 1992, ApJ, 399, 405

Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230
Weigel A. K., Schawinski K., Caplar N., Wong O. I., Treister E., Trakhtenbrot B., 2017, ApJ, 845, 134
Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87
Weinberger R., et al., 2017a, MNRAS, 465, 3291
Weinberger R., Ehlert K., Pfrommer C., Pakmor R., Springel V., 2017b, MNRAS, 470, 4530
Whiley I. M., et al., 2008, MNRAS, 387, 1253
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
White S. D. M., Silk J., 1979, ApJ, 231, 1
Wiersma R. P. C., Schaye J., Smith B. D., 2009a, MNRAS, 393, 99
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, MNRAS, 399, 574
Williams L. L. R., Babul A., Dalcanton J. J., 2004, ApJ, 604, 18
Willman B., Governato F., Wadsley J., Quinn T., 2004, MNRAS, 355, 159
Wright E. L., et al., 2010, Astron. J., 140, 1868
Wu X., Fang L., 1997, ApJ, 483, 62
Wuyts S., Cox T. J., Hayward C. C., Franx M., Hernquist L., Hopkins P. F., Jonsson P., van Dokkum P. G., 2010, ApJ, 722, 1666
Wylezalek D., Zakamska N. L., 2016, MNRAS, 461, 3724
Yang K. H.-Y., Reynolds C. S., 2016a, ApJ, 818, 181
Yang K. H.-Y., Reynolds C. S., 2016b, ApJ, 829, 90
Yoshida N., Stoehr F., Springel V., White S. D. M., 2002, MNRAS, 335, 762
Zahid H. J., Geller M. J., Damjanov I., Sohn J., 2019, preprint (arXiv:1902.04087)
Zakamska N. L., Narayan R., 2003, ApJ, 582, 162
Zavala J., Balogh M. L., Afshordi N., Ro S., 2012, MNRAS, 426, 3464
Zeldovich Y. B., 1970, A&A, 5, 84
Zhang Y.-Y., et al., 2010, ApJ, 711, 1033
Zhang Y. Y., Andernach H., Caretta C. a., Reiprich T. H., Böhringer H., Puchwein E., Sijacki D., Girardi M., 2011, A&A, 526, A105
Zhang Y., et al., 2016, ApJ, 816, 98
Zhang Y., et al., 2018, preprint, pp 1–19 (arXiv:1812.04004)
Zhuravleva I., et al., 2014, Nature, 515, 85
Zibetti S., 2007, Proc. Int. Astron. Union, 3, 176

Zibetti S., White S. D. M., Schneider D. P., Brinkmann J., 2005, MNRAS, 358, 949
Zou S., Maughan B. J., Giles P. A., Vikhlinin A., Pacaud F., Burenin R., Hornstrup A., 2016, MNRAS, 463, 820
Zubeldia I., Challinor A., 2019, preprint (arXiv:1904.07887)
Zwicky F., 1933, Helv. Phys. Acta, 6, 110
Zwicky F., 1937, ApJ, 86, 217
de Haan T., et al., 2016, ApJ, 832, 95
de Jong J. T. A., Verdoes Kleijn G. A., Kuijken K. H., Valentijn E. A., 2013, Exp. Astron., 35, 25
de Plaa J., et al., 2017, A&A, 607, A98
van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649
van der Burg R. F. J., Muzzin A., Hoekstra H., Wilson G., Lidman C., Yee H. K. C., 2014, A&A, 561, A79
von der Linden A., et al., 2014a, MNRAS, 439, 2
von der Linden A., et al., 2014b, MNRAS, 443, 1973